#### PATENT COOPERATIO, REATY

#### From the INTERNATIONAL BUREAU

#### **PCT**

#### **NOTIFICATION OF ELECTION**

(PCT Rule 61.2)

Date of mailing (day/month/year)

**Assistant Commissioner for Patents** United States Patent and Trademark Office Box PCT Washington, D.C.20231

**ETATS-UNIS D'AMERIQUE** 

in its capacity as elected Office

04 July 2000 (04.07.00) International application No. Applicant's or agent's file reference PCT/EP99/07952 HR199-WO International filing date (day/month/year) Priority date (day/month/year) 20 October 1999 (20.10.99) 31 October 1998 (31.10.98) **Applicant** RABENHORST, Jürgen et al

1.	The designated Office is hereby notified of its election made:
	X in the demand filed with the International Preliminary Examining Authority on:
	19 April 2000 (19.04.00)
	in a notice effecting later election filed with the International Bureau on:
2.	The election X was
	was not
	made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO 34, chemin des Colombettes 1211 Geneva 20, Switzerland

Authorized officer

Olivia RANAIVOJAONA

Telephone No.: (41-22) 338.83.38

Facsimile No.: (41-22) 740.14.35

### PATENT COOPERATION TREATY

## **PCT**

## RECEIVED

# INTERNATIONAL PRELIMINARY EXAMINATION REPORT AUG 3 1 2001

(PCT Article 36 and Rule 70)

**TECH CENTER 1600/2900** 

Applicant's or agent's file reference HR199-WO	FOR FURTHER ACTION		ionofTransmittalofInternational Preliminary Report (Form PCT/IPEA/416)	
International application No. PCT/EP99/07952	International filing date (day/n 20 October 1999 (20.		Priority date (day/month/year) 31 October 1998 (31.10.98)	
International Patent Classification (IPC) or n C12N 15/00, 1/00, 9/02, 9/00, 9/		42		
Applicant	HAARMANN & REIME	ER GMBH		
and is transmitted to the applicant acc.  This REPORT consists of a total of  This report is also accompaniamended and are the basis for 70.16 and Section 607 of the	sheets, includir	ng this cover s f the description	ational Preliminary Examining Authority  heet.  on, claims and/or drawings which have been tions made before this Authority (see Rule	
3. This report contains indications relating to the following items:  I Basis of the report  Priority  III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability  IV Lack of unity of invention  V Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement  VI Certain documents cited  VII Certain defects in the international application  VIII Certain observations on the international application				
Date of submission of the demand 19 April 2000 (19.04)		f completion of	of this report anuary 2001 (31.01.2001)	
Name and mailing address of the IPEA/EP	Author	ized officer		
Facsimile No.	Teleph	Telephone No.		

				,
		•	•	
	·•_			
· ·				
	÷			
				,

International application No.

### INTERNATIONAL PRELIMINARY EXAMINATION REPORT

PCT/EP99/07952

I.	I. Basis of the rep rt						
1.	With	regard to	the elements of the international application:*				
	$\Box$	the inte	rnational application as originally filed				
	$\overline{\boxtimes}$	the desc	cription:				
		pages	1-108	, as originally filed			
		pages		, filed with the demand			
		pages	, filed with the letter of				
	X	the clair	ms:	ļ			
	ست	pages	1-16	, as originally filed			
		pages	, as amended (together with	any statement under Article 19			
		pages		, filed with the demand			
		pages	, filed with the letter of				
	$\boxtimes$	the drav	wings:				
		pages	1/3-3/3	, as originally filed			
		pages		, filed with the demand			
		pages	, filed with the letter of				
	$\boxtimes$	the seque	ence listing part of the description:				
		pages	1-75	, as originally filed			
		pages		, filed with the demand			
		pages	, filed with the letter of				
2.	the i	nternation e elemen the lan	to the language, all the elements marked above were available or furnished to this Au nal application was filed, unless otherwise indicated under this item. Its were available or furnished to this Authority in the following language arguage of a translation furnished for the purposes of international search (under Rule 23 arguage of publication of the international application (under Rule 48.3(b)).	which is:			
		the lan or 55.3	nguage of the translation furnished for the purposes of international preliminary example.				
3.	With prelii	minary e. contair filed to furnish furnish The st interna The st been fi	to any nucleotide and/or amino acid sequence disclosed in the international examination was carried out on the basis of the sequence listing:  med in the international application in written form.  regether with the international application in computer readable form.  med subsequently to this Authority in written form.  red subsequently to this Authority in computer readable form.  tatement that the subsequently furnished written sequence listing does not go ational application as filed has been furnished.  The application is identical to the constant of the constant	beyond the disclosure in the			
4.			the claims, Nos the drawings, sheets/fig				
5.		beyond	port has been established as if (some of) the amendments had not been made, since the disclosure as filed, as indicated in the Supplemental Box (Rule 70.2(c)).**				
	in th	is report 70.17).	sheets which have been furnished to the receiving Office in response to an invitation of as "originally filed" and are not annexed to this report since they do not con	ntain amendments (Rule 70.16			
**	Any	replacem	ent sheet containing such amendments must be referred to under item 1 and annexed to	o inis report.			

•

#### INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP 99/07952

Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability;
 citations and explanations supporting such statement

1. Statement			
Novelty (N)	Claims	2, 7-10, 14	YES
	Claims	1, 3-6, 11-13, 15-16	NO
Inventive step (IS)	Claims		YES
	Claims	1-16	NO
Industrial applicability (IA)	Claims	1-16	_ YES
	Claims		NO NO

#### 2. Citations and explanations

This report makes reference to the following document:

D1 EP-A-0 845 532.

#### 1. Novelty

D1 discloses the enzymes of Pseudomonas sp. HR 199, which are responsible for the degradation of eugenol, and its gene sequences (Figure 1; SEQ ID NOS: 1-42). Two different strategies were used for their characterisation. Mutants of this strain were isolated using nitroguanidine and classified with respect to their capacity to be able to use eugenol, ferulic acid and vanillin as a C and energy source (page 7, Example 1). For example, it was demonstrated that the mutant 6164, which was no longer able to use eugenol but, like the wild type, was able to utilise ferulic acid and vanillin, could be complemented by the coniferyl alcohol dehydrogenase (CADH) gene (page 8, Example 5). This gene was further identified by the cleaning of the enzyme CADH, the determination of its Nterminal amino-acid sequence and the formation of a gene probe which could hybridise with the isolated gene (page 9, Example 6).

	(A)	,	

#### INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP 99/07952

The subject matter of Claims 1, 3 to 6, 11 to 13 and 15 to 16 is anticipated by the method disclosed in D1. For this reason, Claims 1, 3 to 6, 11 to 13 and 15 to 16 do not meet the requirements of PCT Article 33(2).

#### 2. Inventive step

The method disclosed in D1 anticipates the subject matter of the present application.

The only difference from D1 is that the genes were deactivated by means of methods in genetic engineering (insertion or deletion) (Claims 2, 7 to 10 and 14) and not by means of chemical mutagenesis. This technical feature is not inventive because it is well known to a person skilled in the art.

Consequently, the subject matter of Claims 2, 7-10 and 14 does not involve an inventive step and therefore does not meet the criterion stipulated in PCT Article 33(3).

		,	
		1	
•			

#### INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/EP 99/07952

#### VIII. Certain observations on the international application

The following observations on the clarity of the claims, description, and drawings or on the question whether the claims are fully supported by the description, are made:

- The present application demonstrates that this method functions with the bacterium Pseudomonas. However, it does not provide any basis or example for a different unicellular organism and, particularly, for a "multicellular organism". For this reason, Claim 1 is not sufficiently supported and thus does not meet the requirements of PCT Article 6.
- Claim 1 clearly has to be restricted to non-human animals.
- 3. It looks as if Claims 7 and 8 should be dependent on Claim 6.
- 4. The expressions "preferably" or "in particular" only introduce a preferred variant. However, this does not limit the scope of Claims 4-5 and 12.
- 5. In Claim 7, the gene structures should not be defined with a reference to the drawing but by technical features.
- 6. The gene structures in Claim 8 should not be defined with reference to the drawing but to the sequence.

	••••	•

## VERTRAG BER DIE INTERNATIONALE USAMMENARBEIT AUF DEMOI **GEBIET DES PATENTWESENS**

WIPO

PCT

## **PCT**

### INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT

(Artikel 36 und Regel 70 PCT)

Aktenzeichen des Anmelders oder Anwalts	T					
HR199-WO	WEITERES VORGEHEN	siehe Mitteilung über die Übersendung des internationalen vorläufigen Prüfungsberichts (Formblatt PCT/IPEA/416)				
Internationales Aktenzeichen	Internationales Anmeldedatum(Tag	g/Monat/Jahr) Prioritätsdatum (Tag/Monat/Tag)				
PCT/EP99/07952	20/10/1999	31/10/1998				
Internationale Patentklassifikation (IPK) oder (C12N15/00	nationale Klassifikation und IPK					
Anmelder						
HAARMANN & REIMER GMBH et a	l					
Dieser internationale vorläufige Prü- Behörde erstellt und wird dem Anme		der internationalen vorläufigen Prüfung beauftragten elt.				
2. Dieser BERICHT umfaßt insgesamt	5 Blätter einschließlich dieses	Deckblatts.				
und/oder Zeichnungen, die geä	ndert wurden und diesem Bericl	sich um Blätter mit Beschreibungen, Ansprüchen nt zugrunde liegen, und/oder Blätter mit vor dieser ınd Abschnitt 607 der Verwaltungsrichtlinien zum PCT).				
Diese Anlagen umfassen insgesam	t Blätter.					
3. Dieser Bericht enthält Angaben zu f	olgenden Punkten:					
I ⊠ Grundlage des Berichts	<b>3</b>					
II □ Priorität		,				
III   Keine Erstellung eines	Gutachtens über Neuheit, erfind	erische Tätigkeit und gewerbliche Anwendbarkeit				
IV	eit der Erfindung					
		der Neuheit, der erfinderischen Tätigkeit und der gen zur Stützung dieser Feststellung				
VI 🔲 Bestimmte angeführte t	Jnterlagen					
VII 🛚 Bestimmte Mängel der	internationalen Anmeldung					
VIII 🛛 Bestimmte Bemerkunge	en zur internationalen Anmeldun	g				
Datum der Einreichung des Antrags	Datum d	er Fertigstellung dieses Berichts				
19/04/2000	31.01.20	001				
Name und Postanschrift der mit der internation Prüfung beauftragten Behörde:	nalen vorläufigen Bevollm	ächtigter Bediensteter				
Europäisches Patentamt D-80298 München Tel. +49 89 2399 - 0 Tx; 523656	Buchet	z, A				
Fax: +49 89 2399 - 4465		+49 89 2399 7401				

		,
•		



### INTERNATIONALER VORLÄUFIGER **PRÜFUNGSBERICHT**

Internationales Aktenzeichen PCT/EP99/07952

I.	Gr	undlage des B ric	hts
1.	Art nic	tikel 14 hin vorgeleg	erstellt auf der Grundlage ( <i>Ersatzblätter, die dem Anmeldeamt auf eine Aufforderung nach</i> It wurden, gelten im Rahmen dieses Berichts als "ursprünglich eingereicht" und sind ihm ie keine Änderungen enthalten.): n:
	1-1	108	ursprüngliche Fassung .
	Pa	tentansprüche, Nr.	. <del>.</del>
	1-1	6	ursprüngliche Fassung
	Zei	ichnungen, Blätter	:
	1/3	-3/3	ursprüngliche Fassung
	Sec	quenzprotokoll in d	der Beschreibung, Seiten:
	1-7	5, in der ursprünglic	ch eingereichten Fassung.
2.	die	internationale Anmo	he: Alle vorstehend genannten Bestandteile standen der Behörde in der Sprache, in der eldung eingereicht worden ist, zur Verfügung oder wurden in dieser eingereicht, sofern hts anderes angegeben ist.
	Die eing	Bestandteile stand gereicht; dabei hand	en der Behörde in der Sprache: zur Verfügung bzw. wurden in dieser Sprache delt es sich um
		die Sprache der Ü Regel 23.1(b)).	bersetzung, die für die Zwecke der internationalen Recherche eingereicht worden ist (nac
		die Veröffentlichun	gssprache der internationalen Anmeldung (nach Regel 48.3(b)).
			bersetzung, die für die Zwecke der internationalen vorläufigen Prüfung eingereicht worden
3.	Hin: inte	sichtlich der in der i rnationale vorläufig	nternationalen Anmeldung offenbarten <b>Nucleotid- und/oder Aminosäuresequenz</b> ist die e Prüfung auf der Grundlage des Sequenzprotokolls durchgeführt worden, das:
	×	in der international	en Anmeldung in schriftlicher Form enthalten ist.
	$\boxtimes$		internationalen Anmeldung in computerlesbarer Form eingereicht worden ist.
			achträglich in schriftlicher Form eingereicht worden ist.
		bei der Behörde na	achträglich in computerlesbarer Form eingereicht worden ist.

 $\square$  Die Erklärung, daß das nachträglich eingereichte schriftliche Sequenzprotokoll nicht über den

☐ Die Erklärung, daß die in computerlesbarer Form erfassten Informationen dem schriftlichen

Offenbarungsgehalt der internationalen Anmeldung im Anmeldezeitpunkt hinausgeht, wurde vorgelegt.





# INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT

Internationales Aktenzeichen PCT/EP99/07952

Sequenzprotokoll entsprechen, wurde vorgelegt.

		, ,	•				
4.	Auf	grund der Änderunger	n sind folgend	le Uı	nterlagen fort	gefallen:	
		Beschreibung,	Seiten:				
		Ansprüche,	Nr.:				
		Zeichnungen,	Blatt:				
5.			en nach Auffa	assu	ng der Behör	en) der Änderungen erstellt worden, da diese aus den de über den Offenbarungsgehalt in der ursprünglich	
	(Auf Ersatzblätter, die solche Änderungen enthalten, ist unter Punkt 1 hinzuweisen;sie sind diesem Berich beizufügen).						
6.	Etw	aige zusätzliche Bem	erkungen:				
V.						ich der Neuheit, der erfinderischen Tätigkeit und der rungen zur Stützung dieser Feststellung	
1.	Fes	tstellung					
	Neu	uheit (N)	_	a: lein:	Ansprüche Ansprüche	2, 7-10, 14 1, 3-6, 11-13, 15-16	
	Erfi	nderische Tätigkeit (E	•	a: lein:	Ansprüche Ansprüche	1-16	
	Gev	verbliche Anwendbark		a: lein:	Ansprüche Ansprüche	1-16	

2. Unterlagen und Erklärungen siehe Beiblatt

#### VIII. Bestimmte Bemerkungen zur internationalen Anmeldung

Zur Klarheit der Patentansprüche, der Beschreibung und der Zeichnungen oder zu der Frage, ob die Ansprüche in vollem Umfang durch die Beschreibung gestützt werden, ist folgendes zu bemerken: siehe Beiblatt





## INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT - BEIBLATT

Internationales Aktenzeichen PCT/EP99/07952

#### Zu Punkt V

Begründete Feststellung nach Artikel 35(2) hinsichtlich der Neuheit, der erfinderischen Tätigkeit und der gewerblichen Anwendbarkeit; Unterlagen und Erklärungen zur Stützung dieser Feststellung

Es wird auf das folgende Dokument verwiesen:

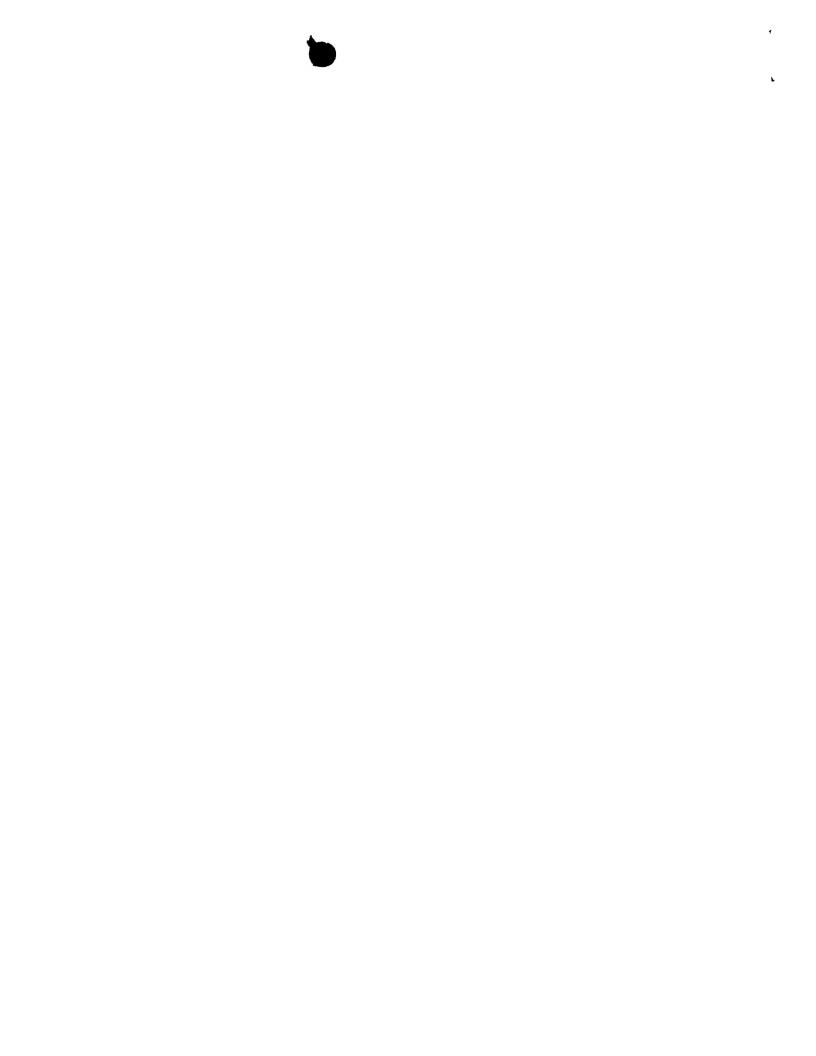
D1: EP 0 845 532 A

#### 1) Neuheit:

- D1 offenbart die Enzyme von *Pseudomonas* sp. HR 199, die verantwortlich für den Abbau von Eugenol sind, und ihre Gensequenzen (Fig. 1; SEQ ID NOS: 1-42). Zwei verschiedene Strategien wurden für ihre Charakterisierung benutzt. Entweder wurden Mutanten dieses Stammes mittels Nitroguanidin isoliert und bezüglich ihres Vermögens, Eugenol, Ferulasäure und Vanillin als C- und Energiequelle nutzen zu können, klassifiziert (S. 7, Beispiel 1). Zum Beispiel wurde gezeigt, daß die Mutante 6164, die nicht mehr in der Lage war, Eugenol zu nutzen aber wie der Wildtyp, Ferulasäure und Vanillin zu verwerten vermochte, mit dem Coniferylalkohol-Dehydrogenase (CADH) Gen komplementiert werden konnte (S. 8, Beispiel 5). Dieses Gen wurde weiter identifiziert durch die Reinigung des Enzymes CADH, die Bestimmung seiner N-terminalen Aminosäuresequenz und die Ausbildung einer Gensonde, die mit dem isolierten Gen hybridisieren konnte (S. 9, Beispiel 6).
- Der Gegenstand der Ansprüche 1, 3-6, 11-13 und 15-16 wird durch das in D1 offenbarte Verfahren vorweggenommen. Aus diesem Grund erfüllen die Ansprüche 1, 3-6, 11-13 und 15-16 nicht die Erfordernisse des Artikels 33(2) PCT.

#### 2) Erfinderische Tätigkeit:

- Das in D1 offenbarte Verfahren nimmt den Gegenstand der vorliegenden Anmeldung vorweg.





## INTERNATIONALER VORLÄUFIGER PRÜFUNGSBERICHT - BEIBLATT

Internationales Aktenzeichen PCT/EP99/07952

- Der einzige Unterschied zu D1 ist, daß die Gene mittels gentechnischer Methoden (Insertion oder Deletion) inaktiviert wurden (Ansprüche 2, 7-10 und 14), und nicht mittels chemischer Mutagenese. Dieses technische Merkmal ist nicht erfinderisch, weil es dem Fachmann wohlbekannt ist.
- Der Gegenstand der Ansprüche 2, 7-10 und 14 beruht somit nicht auf einer erfinderischen Tätigkeit und erfüllt damit nicht das in Artikel 33(3) PCT genannte Kriterium.

#### Zu Punkt VIII

#### Bestimmte Bemerkungen zur internationalen Anmeldung

- 1) Die vorliegende Anmeldung zeigt, daß dieses Verfahren mit dem Bakterium *Pseudomonas* funktioniert. Es gibt aber keine Grundlage oder kein Beispiel für einen anderen einzelligen Organismus und, um so mehr, für einen "mehrzelligen Organismus". Aus diesem Grund ist Anspruch 1 nicht ausreichend gestützt und erfüllt somit nicht die Erfordernisse des Artikels 6 PCT.
- 2) Anspruch 1 muß ausdrücklich auf nicht-menschliche Tiere beschränkt werden.
- 3) Es sieht so aus, daß Ansprüche 7 und 8 von Anspruch 6 abhängen sollten.
- 4) Das Wort "vorzugsweise" oder "insbesondere" führt nur eine bevorzugte Variante ein. Das limitiert den Umgang der Ansprüche 4-5 und 12 aber nicht.
- 5) In Anspruch 7 sollten die Gen-Strukturen nicht mit einer Referenz zur Abbildung sondern mit technischen Merkmalen definiert werden.
- 6) Die Gen-Strukturen in Anspruch 8 sollten nicht mit Referenz zur Abbildung sondern zur Sequenz definiert werden.



itional Application No

PCT/EP 99/07952 CLASSIFICATION OF SUBJECT MATTER C 7 C12N15/00 C12N C12N9/00 C12N9/88C12N1/00 C12N9/02 C12P7/22 C12P7/24 C12P7/42 C12N9/10 According to International Patent Classification (IPC) or to both national classification and IPC B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) C12N C12P IPC 7 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Relevant to claim No. Citation of document, with indication, where appropriate, of the relevant passages Category 1,3-6,EP 0 845 532 A (HAARMANN & REIMER GMBH) χ 3 June 1998 (1998-06-03) 11-16 cited in the application the whole document WO 97 35999 A (INST OF FOOD RESEARCH 1,4,5, χ :NARBAD ARJAN (GB); RHODES MICHAEL JOHN 12-15 CHAR) 2 October 1997 (1997-10-02) cited in the application the whole document -/--Further documents are listed in the continuation of box C. Patent family members are listed in annex. Special categories of cited documents : "T" later document published after the international filing date or priority date and not in conflict with the application but called to understand the principle or theory underlying the "A" document defining the general state of the lart which is not considered to be of particular relevance invention "E" earlier document but published on or after the international "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other, such document is combined with one or more other. citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or ments, such combination being obvious to a person skilled other means document published prior to the international filing date but later than the priority date claimed "&" document member of the same patent family Date of mailing of the international search report Date of the actual completion of the international search 26/06/2000 7 June 2000 Authorized officer Name and mailing address of the ISA European Patent Office. P.B. 5818 Patentiaan 2

1

NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl. Fax: (+31-70) 340-3016

Oderwald, H

## INTERNATIONAL ARCH REPORT

Int. Itional April cation No PCT/EP 99/07952

	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
ategory *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No
(	GASSON M J ET AL.: "Metabolism of ferulic acid to vanillin" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 7, 13 February 1998 (1998-02-13), pages 4163-4170, XP002139585 cited in the application the whole document	1,4,5, 12-15
(	DATABASE WPI Section Ch, Week 199340 Derwent Publications Ltd., London, GB; Class B05, AN 1993-316614 XP002139586 & JP 05 227980 A (TAKASAGO PERFUMERY CO LTD), 7 September 1993 (1993-09-07) abstract	1,4,5, 12,15
A	EP 0 583 687 A (HAARMANN & REIMER GMBH) 23 February 1994 (1994-02-23) cited in the application the whole document	

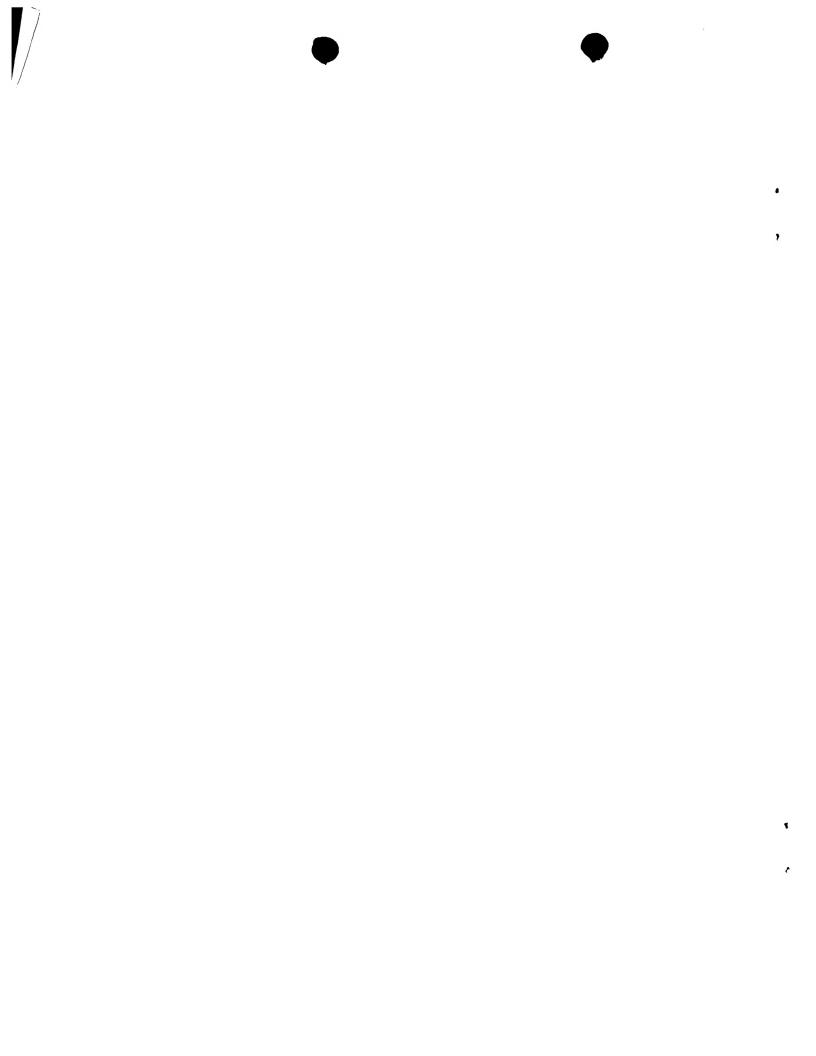
1

## IN NATIONAL SEARCH REPORT

information on patent family members

In. ational Application No PCT/EP 99/07952

	atent document d in search report		Publication gate	Patent family member(s)	Publication date
EP	0845532	A	03-06-1998	DE 19649655 A JP 10155496 A	04-06-1998 16-06-1998
WO	9735999	A	02-10-1997	AU 2038597 A CA 2250043 A EP 0904396 A	17-10-1997 02-10-1997 31-03-1999
JP	5227980	Α	07-09-1993	NONE	
EP	0583687	A	23-02-1994	DE 4227076 A DE 59309293 D JP 6153924 A US 5371013 A US 5510252 A US 5976863 A	24-02-1994 25-02-1999 03-06-1994 06-12-1994 23-04-1996 02-11-1999



#### LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
ΑU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland		Republik Mazedonien	TR	Türkei
BG	Bulgarien	HU	Ungarn	ML	Mali	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MN	Mongolei	UA	Ukraine
BR	Brasilien	IL	Israel	MR	Mauretanien	UG	Uganda
BY	Belarus	18	Island	MW	Malawi	US	Vereinigte Staaten von
CA	Kanada	IT	Italien	MX	Mexiko		Amerika
CF	Zentralafrikanische Republik	JP	Japan	NE	Niger	UZ	Usbekistan
CG	Kongo	KE	Kenia	NL	Niederlande	VN	Vietnam
CH	Schweiz	KG	Kirgisistan	NO	Norwegen	YU	Jugoslawien
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik	NZ	Neusceland	zw	Zimbabwe
CM	Kamerun		Korea	PL	Polen		
CN	China	KR	Republik Korea	PT	Portugal		
CU	Kuba	KZ	Kasachstan	RO	Rumänien		
CZ	Tschechische Republik	LC	St. Lucia	RU	Russische Föderation		
DE	Deutschland	LI	Liechtenstein	SD	Sudan		
DK	Dänemark	LK	Sri Lanka	SE	Schweden		
EE	Estland	LR	Liberia	SG	Singapur		



### WELTORGANISATION FÜR GEISTIGES E-GENTUM

Internationales Büro

INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 7:

C12N 15/00, 1/00, 9/02, 9/00, 9/88, 9/10, C12P 7/22, 7/24, 7/42

(11) Internationale Veröffentlichungsnummer: **A3** 

(43) Internationales

Veröffentlichungsdatum:

11. Mai 2000 (11.05.00)

WO 00/26355

(21) Internationales Aktenzeichen:

PCT/EP99/07952

(22) Internationales Anmeldedatum: 20. Oktober 1999 (20.10.99)

(30) Prioritätsdaten:

198 50 242.7

31. Oktober 1998 (31.10.98)

DE

(71) Anmelder (für alle Bestimmungsstaaten ausser US): HAAR-MANN & REIMER GMBH [DE/DE]; D-37601 Holzminden (DE).

(72) Erfinder; und

- (75) Erfinder/Anmelder (nur für US): RABENHORST, Jürgen Weserblick 13, D-37671 Höxter (DE). [DE/DE]; STEINBÜCHEL, Alexander [DE/DE]; Rönnenthal 27, D-48341 Altenberge (DE). PRIEFERT, Horst [DE/DE]; Potthoffskamp 5, D-48291 Telgte (DE). OVERHAGE, Jörg [DE/DE]; Rotherbachstrasse 42 a, D-59132 Bergkamen (DE).
- (74) Anwalt: MANN, Volker, Bayer Aktiengesellschaft, D-51368 Leverkusen (DE).

(81) Bestimmungsstaaten: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Veröffentlicht

Mit internationalem Recherchenbericht.

(88) Veröffentlichungsdatum des internationalen Recherchenbe-9. November 2000 (09.11.00) richts:

- (54) Title: CONSTRUCTION OF PRODUCTION STRAINS FOR PRODUCING SUBSTITUTED PHENOLS BY SPECIFICALLY INACTIVATING GENES OF THE EUGENOL AND FERULIC ACID CATABOLISM
- (54) Bezeichnung: KONSTRUKTION VON PRODUKTIONSSTÄMMEN FÜR DIE HERSTELLUNG VON SUBSTITUIERTEN PHENOLEN DURCH GEZIELTE INAKTIVIERUNGEN VON GENEN DES EUGENOL- UND FERU-LASÄURE-KATABOLISMUS

#### (57) Abstract

The invention relates to a transformed and/or mutagenated unicellular or multicellular organism which is characterized in that enzymes of the eugenol and/or ferulic acid catabolism are deactivated in such a manner that the intermediates coniferyl alcohol, coniferyl aldehyde, ferulic acid, vanillin and/or vanillinic acid are accumulated.

#### (57) Zusammenfassung

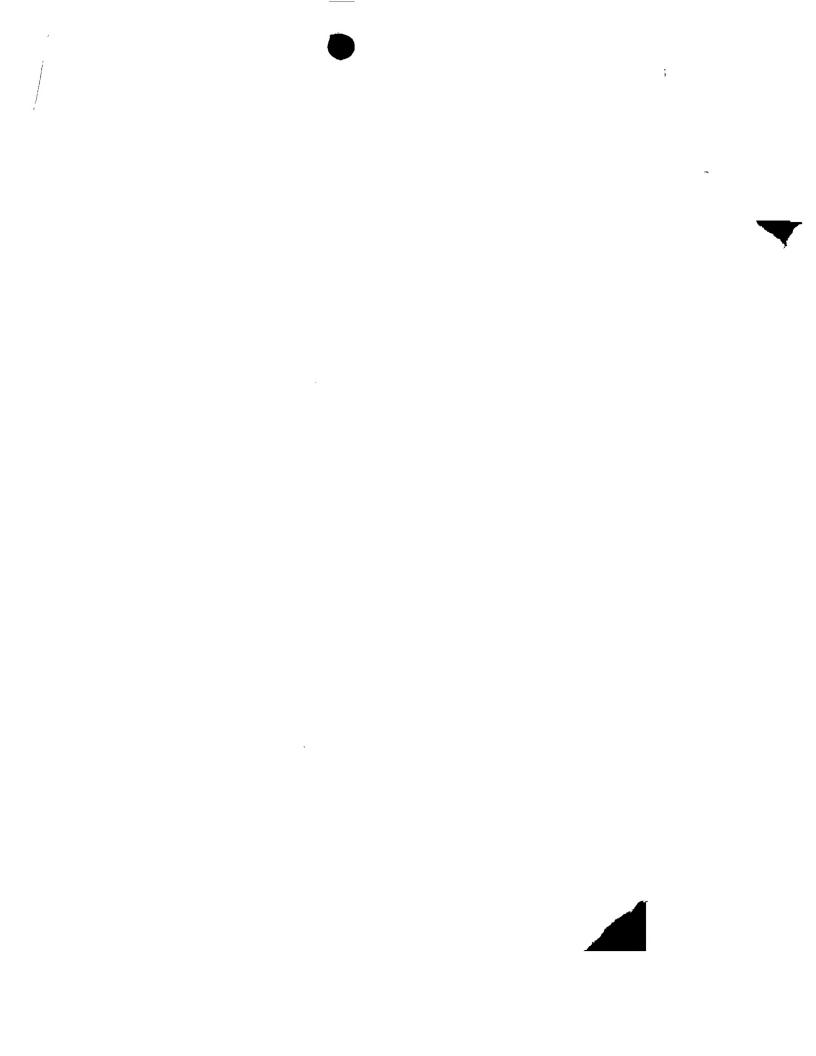
Die vorliegende Erfindung betrifft einen transformierten und/oder mutagenisierten ein- oder mehrzelligen Organismus, der dadurch gekennzeichnet ist, daß Enzyme des Eugenol- und/oder Ferulasäure-Katabolismus derart inaktiviert sind, daß eine Akkumulation der Intermediate Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure erfolgt,

## TION ER RECHERCHENBERICHT

zu Veröffentlichungen, die zur selben Patentfamilie genoren

in. Includes Aktenzeichen PCT/EP 99/07952

rchenbericht Patentdokument		Datum der Veröffentlichung		litglied(er) der Patentfamilie	Datum der Veröffentlichung
0845532	Α	03-06-1998	DE JP	19649655 A 10155496 A	04-06-1998 16-06-1998
9735999	Α	02-10-1997	AU CA EP	2038597 A 2250043 A 0904396 A	17-10-1997 02-10-1997 31-03-1999
5227980	Α	07-09-1993	KEI	VE	
0583687	A	23-02-1994	DE DE JP US US	4227076 A 59309293 D 6153924 A 5371013 A 5510252 A 5976863 A	24-02-1994 25-02-1999 03-06-1994 06-12-1994 23-04-1996 02-11-1999



## INTERNATIONALER RECHE HENBERICHT

In: ationales Akenzeichen PCT/EP 99/07952

		CI/EP 99/07952	
C.(Fortsetz	ung) ALS WESENTLICH ANGESEHENE UNTERLAGEN		
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommend	en Teile Betr. Anspruch	Nr.
X	GASSON M J ET AL.: "Metabolism of ferulic acid to vanillin" JOURNAL OF BIOLOGICAL CHEMISTRY, Bd. 273, Nr. 7, 13. Februar 1998 (1998-02-13), Seiten 4163-4170, XP002139585 in der Anmeldung erwähnt das ganze Dokument	1,4, 12-1	
X	DATABASE WPI Section Ch, Week 199340 Derwent Publications Ltd., London, GB; Class B05, AN 1993-316614 XP002139586 & JP 05 227980 A (TAKASAGO PERFUMERY CO LTD), 7. September 1993 (1993-09-07) Zusammenfassung	1,4,5	
A	EP 0 583 687 Å (HAARMANN & REIMER GMBH) 23. Februar 1994 (1994-02-23) in der Anmeldung erwähnt das ganze Dokument		

₄tionales Aktenzeichen PCT/EP 99/07952

a. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES IPK 7 C12N15/00 C12N1/00 C12N9/02 C12N9/00 C12N9/88 C12N9/10 C12P7/22 C12P7/24 C12P7/42

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

#### B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)

IPK 7 C12N C12P

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

#### C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
X	EP 0 845 532 A (HAARMANN & REIMER GMBH) 3. Juni 1998 (1998-06-03) in der Anmeldung erwähnt das ganze Dokument	1,3-6, 11-16
X	WO 97 35999 A (INST OF FOOD RESEARCH; NARBAD ARJAN (GB); RHODES MICHAEL JOHN CHAR) 2. Oktober 1997 (1997-10-02) in der Anmeldung erwähnt das ganze Dokument	1,4,5, 12-15

Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen	X Siehe Anhang Patentfamilie
Besondere Kategorien von angegebenen Veröffentlichungen :	"T" Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum
"A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist	Anmeldung nicht kollidiert, sondern nur zum. Verständnis des der
"E" älteres Dokument, das jedoch erst am oder inach dem internationalen Anmeldedatum veröffentlicht worden ist	Enindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist
"L" Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft er- scheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung beledt werden	erfinderischer Tätigkot berühend betreitet den Alle in der der auf

soll oder die aus einem anderen besonderen Grund angegeben ist (wie Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung Verorientlichung von besonderer Bedeutung; die beanspruchte Erfindu kann nicht als auf erlinderischer Tätigkeit berühend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategone in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist ausgeführt) Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

"&" Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der internationalen Recherche

7. Juni 2000 26/06/2000

Name und Postanschrift der Internationalen Recherchenbehörde Europäisches Patentamt, P.B. 5818 Patentlaan 2

NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Oderwald, H

Formblatt PCT/ISA/210 (Blatt 2) (Juli 1992)

1

Absendedatum des internationalen Recherchenberichts

## PCT

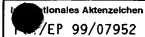
#### INTERNATIONALER RECHERCHENBERICHT

(Artikel 18 sowie Regeln 43 und 44 PCT)

Aktenzeichen des Anmelders oder Anwalts	WEITERES siehe Mitteilung über die Übermittlung des internationalen							
HR199-WO		richts (Formblatt PCT/ISA/220) sowie, soweit chstehender Punkt 5						
Internationales Aktenzeichen	Internationales Anmeldedatum	(Frühestes) Prioritätsdatum (Tag/Monat/Jahr)						
PCT/EP 99/07952	(Tag/Monat/Jahr) 20/10/1999	31/10/1998						
Anmelder	20/10/1///	31/10/1/20						
,								
HAARMANN & REIMER GMBH et a	ı1.							
Dieser internationale Recherchenbericht wurd Artikel 18 übermittelt. Eine Kopie wird dem Int		ehörde erstellt und wird dem Anmelder gemäß						
	_							
Dieser internationale Recherchenbericht umfa		ter. nannten Unterlagen zum Stand der Technik bei.						
X Darüber hinaus liegt ihm jew	ens enne kophe der in diesem bencht ge	mannen Offenagen zum Gland der Technik bei.						
Grundlage des Berlchts								
	rnationale Recherche auf der Grundlage ereicht wurde, sofern unter diesem Punk	der internationalen Anmeldung in der Sprache d nichts anderes angegeben ist.						
Die internationale Recherch Anmeldung (Regel 23.1 b)) (	e ist auf der Grundlage einer bei der Beh durchgeführt worden.	nörde eingereichten Übersetzung der internationalen						
b. Hinsichtlich der in der internationale	n Anmeldung offenbarten <b>Nucleotid– ur</b> equenzprotokolls durchgeführt worden, d	nd/oder Aminosäuresequenz ist die internationale						
I ~~~	dung in Schriflicher Form enthalten ist.	uas -						
zusammen mit der internation	onalen Anmeldung in computerlesbarer F	Form eingereicht worden ist.						
bei der Behörde nachträglich	n in schriftlicher Form eingereicht worder	n ist.						
_ <u> </u>	n in computerlesbarer Form eingereicht v							
Die Erklärung, daß das nach internationalen Anmeldung i	nträglich eingereichte schriftliche Sequen m Anmeldezeitpunkt hinausgeht, wurde	nzprotokoll nicht über den Offenbarungsgehalt der vorgelegt.						
Die Erklärung, daß die in co wurde vorgelegt.	mputerlesbarer Form erfaßten Informatio	onen dem schriftlichen Sequenzprotokoll entsprechen,						
2. Bestimmte Ansprüche hat	en sich als nicht recherchierbar erwie	esen (siehe Feld I).						
3. Mangeinde Einhettlichkett	der Erfindung (siehe Feld II).							
4. Hinsichtlich der Bezelchnung der Erfin	dung							
wird der vom Anmelder eing	ereichte Wortlaut genehmigt.							
wurde der Wortlaut von der	Behörde wie folgt festgesetzt:							
5. Hinsichtlich der <b>Zusammenfassung</b>	5. Hinsichtlich der <b>Zusammenfassung</b>							
wurde der Wortlaut nach Re	innerhalb eines Monats nach dem Datu	n Fassung von der Behörde festgesetzt. Der m der Absendung dieses internationalen						
6. Folgende Abbildung der <b>Zelchnungen</b> i	st mit der Zusammenfassung zu veröffer							
wie vom Anmelder vorgesch		keine der Abb.						
1 =	ne Abbildung vorgeschlagen hat.							
weil diese Abbildung die Erfi	indung besser kennzeichnet.							

			-
04			

#### INTERNATIONALER RECHERCHENBERICHT



a. Klassifizierung des anmeldungsgegenstandes IPK 7 C12N15/00 C12N1/00 C12N9/02 C12N9/88 C12N9/00 C12N9/10 C12P7/22 C12P7/24 C12P7/42 Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK **B. RECHERCHIERTE GEBIETE** Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole) IPK 7 C12N C12P Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe) C. ALS WESENTLICH ANGESEHENE UNTERLAGEN Kategorie° Bezeigenung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile Betr. Anspruch Nr. X EP 0 845 532 A (HAARMANN & REIMER GMBH) 1,3-6, 3./Juni 1998 (1998-06-03) 11-16 j⁄n der Anmeldung erwähnt das ganze Dokument X WO 97 35999 A (INST OF FOOD RESEARCH 1,4,5, ; NARBAD ARJAN (GB); RHODES MICHAEL JOHN 12 - 15CHAR) 2. Oktober 1997 (1997-10-02) in der Anmeldung erwähnt das ganze Dokument -/--Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu χ Siehe Anhang Patentfamilie X ° Besondere Kategorien von angegebenen Veröffentlichungen Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondem nur zum Verständnis des der "A" Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist Erfindung zugrundellegenden Prinzips oder der ihr zugrundellegenden Theorie angegeben ist "E" älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist "X" Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden "L\* Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft er-scheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung soll oder die aus einem anderen besonderen Grund angegeben ist (wie kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet ausgeführt) werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist "O" Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht "P" Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist \*& Veröffentlichung, die Mitglied derselben Patentfamilie ist Datum des Abschlusses der internationalen Recherche Absendedatum des internationalen Recherchenberichts 7. Juni 2000 26/06/2000 Bevollmächtigter Bediensteter Name und Postanschrift der Internationalen Recherchenbehörde Europäisches Patentamt, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Oderwald, H Fax: (+31-70) 340-3016

1

## INTERNATIONALER RECHERCHENBERICHT



		EF 99/0/952
C.(Fortsetzu	ng) ALS WESENTLICH ANGESEHENE UNTERLAGEN	
Kategorie°	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht komme	enden Teile Betr. Anspruch Nr.
×	GASSON M J ET AL.: "Metabolism of ferulic acid to vanillin" JOURNAL OF BIOLOGICAL CHEMISTRY, Bd. 273, Nr. 7, 13. Februar 1998 (1998-02-13), Seiten 4163-4170, XP002139585 in der Anmeldung erwähnt das ganze Dokument	1,4,5, 12-15
×	DATABASE WPI Section Ch, Week 199340 Derwent Publications Ltd., London, GB; Class B05, AN 1993-316614 XP002139586 & JP 05 227980 A (TAKASAGO PERFUMERY CO LTD), 7. September 1993 (1993-09-07) Zusammenfassung	1,4,5, 12,15
A V	EP 0 583 687 A (HAARMANN & REIMER GMBH) 23. Februar 1994 (1994-02-23) in der Anmeldung erwähnt das ganze Dokument	
		·

#### INTERNATIONAL SEARCH REPORT

nto on patent family members

EP 99/07952

	tent document in search repor	t	Publication date	Patent family member(s)	Publication date
EP	0845532	Α	03-06-1998	DE 19649655 A JP 10155496 A	04-06-1998 16-06-1998
WO	9735999	Α	02-10-1997	AU 2038597 A CA 2250043 A EP 0904396 A	17-10-1997 02-10-1997 31-03-1999
JP	5227980	Α	07-09-1993	NONE	
EP	0583687	A	23-02-1994	DE 4227076 A DE 59309293 D JP 6153924 A US 5371013 A US 5510252 A US 5976863 A	24-02-1994 25-02-1999 03-06-1994 06-12-1994 23-04-1996 02-11-1999

#### WELTORGANISATION FÜR GEISTIGES EIGENTUM Internationales Büro

INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 7:

C12N 15/00

(21) Internationale Patentklassifikation 7:

(22) (13) Internationale Veröffentlichungsnummer: WO 00/26355

(23) Internationales

Veröffentlichungsdatum: 11. Mai 2000 (11.05.00)

(21) Internationales Aktenzeichen:

PCT/EP99/07952

(22) Internationales Anmeldedatum: 20. Oktober 1999 (20.10.99)

(30) Prioritätsdaten:

198 50 242.7

31. Oktober 1998 (31.10.98) DE

(71) Anmelder (für alle Bestimmungsstaaten ausser US): HAAR-MANN & REIMER GMBH [DE/DE]; D-37601 Holzminden (DE).

(72) Erfinder; und

- (75) Erfinder/Anmelder (nur für US): RABENHORST, Jürgen [DE/DE]; Weserblick 13, D-37671 Höxter (DE). STEINBÜCHEL, Alexander [DE/DE]; Rönnenthal 27, D-48341 Altenberge (DE). PRIEFERT, Horst [DE/DE]; Potthoffskamp 5, D-48291 Telgte (DE). OVERHAGE, Jörg [DE/DE]; Rotherbachstrasse 42 a, D-59132 Bergkamen (DE).
- (74) Anwalt: MANN, Volker, Bayer Aktiengesellschaft, D-51368 Leverkusen (DE).

(81) Bestimmungsstaaten: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, LTZ, VN, VIL, ZA, ZW, ARIPO Patent (GH, GM, KF, LS)

SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, SD, UZ, VN, YU, ZA, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Veröffentlicht

Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.

- (54) Title: CONSTRUCTION OF PRODUCTION STRAINS FOR PRODUCING SUBSTITUTED PHENOLS BY SPECIFICALLY INACTIVATING GENES OF THE EUGENOL AND FERULIC ACID CATABOLISM
- (54) Bezeichnung: KONSTRUKTION VON PRODUKTIONSSTÄMMEN FÜR DIE HERSTELLUNG VON SUBSTITUIERTEN PHENOLEN DURCH GEZIELTE INAKTIVIERUNGEN VON GENEN DES EUGENOL- UND FERULASÄURE-KATABOLISMUS

#### (57) Abstract

The invention relates to a transformed and/or mutagenated unicellular or multicellular organism which is characterized in that enzymes of the eugenol and/or ferulic acid catabolism are deactivated in such a manner that the intermediates coniferyl alcohol, coniferyl aldehyde, ferulic acid, vanillin and/or vanillinic acid are accumulated.

#### (57) Zusammenfassung

Die vorliegende Erfindung betrifft einen transformierten und/oder mutagenisierten ein- oder mehrzelligen Organismus, der dadurch gekennzeichnet ist, daß Enzyme des Eugenol- und/oder Ferulasäure-Katabolismus derart inaktiviert sind, daß eine Akkumulation der Intermediate Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure erfolgt.

HR 199

#### LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss den PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
ΑZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
ВВ	Barbados	GH	Ghana	MG	Madagaskar	TJ	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland		Republik Mazedonien	TR	Türkei
BG	Bulgarien	HU	Ungam	ML	Mali	TT	Trinidad und Tobago
BJ	Benin	IE	Irland	MN	Mongolei	UA	Ukraine
BR	Brasilien	IL	Israel	MR	Mauretanien	UG	Uganda
ΒY	Belarus	18	Island	MW	Malawi	US	Vereinigte Staaten von
CA	Kanada	IT	Italien	MX	Mexiko		Amerika
CF	Zentralafrikanische Republik	JP	Japan	NE	Niger	UZ	Usbekistan
CG	Kongo	KE	Kenia	NL	Niederlande	VN	Vietnam
СН	Schweiz	KG	Kirgisistan	NO	Norwegen	YU	Jugosławien
CI	Cite d'Ivoire	KP	Demokratische Volksrepublik	NZ	Neusceland	zw	Zimbabwe
CM	Kamerun		Korea	PL	Polen		
CN	China	KR	Republik Korea	PT	Portugal		
CU	Kuba	KZ	Kasachstan	RO	Rumanien		
C7.	Tschechische Republik	LC	St. Lucia	RU	Russische Föderation		
DE	Deutschland	LI	Liechtenstein	SD	Sudan		
DK	Dänemark	LK	Sri Lanka	SE	Schweden		
EE	Estland	L.R	Liberia	SG	Singapur		••

Konstruktion von Produktionsstämmen für die Herstellung von substituierten Phenolen durch gezielte Inaktivierungen von Genen des Eugenol- und Ferulasäure-Katabolismus

5

Die vorliegende Erfindung betrifft die Konstruktion von Produktionsstämmen und ein Verfahren für die Herstellung substituierter Methoxyphenole, insbesondere Vanillin.

10

Die DE-A 4 227 076 (Verfahren zur Herstellung substituierter Methoxyphenole und dafür geeigneter Mikroorganismus) beschreibt die Herstellung substituierter Methoxyphenole mit einer neuen *Pseudomonas* sp.. Ausgangsmaterial ist hier Eugenol und die Produkte sind Ferulasäure, Vanillinsäure, Coniferylalkohol und Coniferylaldehyd.

15

Ebenfalls 1995 erscheint ein umfangreiches Review über die Biotransformationsmöglichkeiten mit Ferulasäure von Rosazza et al. (Biocatalytic transformation of ferulic acid: an abundant aromatic natural product; J. Ind. Microbiol. 15:457-471).

20

Die Gene und Enzyme zur Synthese von Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und Vanillinsäure aus *Pseudomonas* sp. wurden in EP-A 0 845 532 beschrieben.

25

30

Die Enzyme zur Umsetzung von trans-Ferulasäure zu trans-Feruloyl-SCoA Ester und weiter zum Vanillin, sowie das Gen für die Spaltung des Esters wurden vom Institute of Food Research, Norwich, GB, in WO 97/35999 beschrieben. 1998 erscheint der Inhalt des Patents auch als wissenschaftliche Publikationen (Gasson et al. 1998. Metabolism of ferulic acid to vanillin. J. Biol. Chem. 273:4163-4170; Narbad and Gasson 1998. Metabolism of ferulic acid via vanillin using a novel CoAdependent pathway in a newly isolated strain of *Pseudomonas fluorescens*. Microbiology 144:1397 - 1405).

Die DE-A 195 32 317 beschreibt die fermentative Gewinnung von Vanillin aus Ferulasäure mit *Amycolatopsis* sp. in hohen Ausbeuten.

Die bekannten Verfahren haben den Nachteil, daß entweder nur sehr geringe Ausbeuten an Vanillin erzielt werden, oder von relativ teuren Edukten ausgegangen wird. Bei dem letztgenannten Verfahren (DE-A 195 32 317) werden zwar hohe Ausbeuten erzielt, jedoch bedingt der Einsatz von *Pseudomonas* sp. HR199 und *Amycolatopsis* sp. HR167 für die Biotransformation von Eugenol zu Vanillin eine zweistufige Fermentationsführung und somit einen erheblichen Kosten- und Zeitaufwand.

Aufgabe der vorliegenden Erfindung ist es daher, Organismen zu konstruieren, die in der Lage sind den preiswerten Rohstoff Eugenol in einem einstufigen Prozeß zu Vanillin umzusetzen.

15

Diese Aufgabe wird durch die Konstruktion von Produktionsstämmen ein- oder mehrzelliger Organismen gelöst, die dadurch gekennzeichnet sind, daß Enzyme des Eugenol- und/oder Ferulasäure-Katabolismus derart inaktiviert sind, daß eine Akkumulation der Intermediate Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure erfolgt.

20

25

30

Der Produktionsstamm kann einzellig oder mehrzellig sein. Demgemäß können Gegenstand der Erfindung Mikroorganismen, Pflanzen oder Tiere sein. Darüber hinaus können auch Extrakte die aus dem Produktionsstamm gewonnen werden zum Einsatz kommen. Erfindungsgemäß werden vorzugsweise einzellige Organismen eingesetzt. Hierbei kann es sich um Mikroorganismen, tierische oder pflanzliche Zellen handeln. Besonders bevorzugt ist erfindungsgemäß der Einsatz von Pilzen und Bakterien. Höchst bevorzugt sind Bakterienarten. Unter den Bakterien können insbesondere Rhodococcus-, Pseudomonas- und Escherichia-Arten nach Veränderung des Eugenol- und/oder Ferulasäure-Katabolismus zum Einsatz kommen.

Die Gewinnung der erfindungsgemäß einsetzbaren Organismen kann im einfachsten Fall mittels bekannter, konventioneller mikrobiologischer Methoden erfolgen. So kann die Enzymaktivität der am Eugenol- und/oder Ferulasäure-Katabolismus beteiligten Proteine durch den Einsatz von Enzym-Hemmstoffen verändert werden. Darüber hinaus kann die Enzymaktivität der am Eugenol- und/oder Ferulasäure-Katabolismus beteiligten Proteine durch Mutation der für diese Proteine kodierenden Gene verändert werden. Derartige Mutationen können nach klassischen Methoden ungerichtet erzeugt werden, wie beispielsweise durch UV-Bestrahlung oder mutationsauslösende Chemikalien.

10

15

20

25

30

5

Ebenso sind gentechnische Methoden zur Gewinnung der erfindungsgemäßen Organismen geeignet, wie Deletionen, Insertionen und/oder Nukleotid-Austausche. So können beispielsweise die Gene der Organismen mit Hilfe von anderen DNA-Elementen (Ω-Elemente) inaktiviert werden. Ebenso können mittels geeigneter Vektoren Austausche der intakten Gene gegen veränderte und/oder inaktivierte Gen-Strukturen durchgeführt werden. Die zu inaktivierenden Gene und die für die Inaktivierung eingesetzten DNA-Elemente können dabei durch klassische Klonierungstechniken oder durch Polymerase-Kettenreaktionen (PCR) gewonnen werden.

In einer möglichen Ausgestaltung der Erfindung kann beispielsweise der Eugenolsowie der Ferulasäure-Katabolismus durch Ω-Element-Insertion oder Einführen von Deletionen in entsprechende Gene verändert werden. Hierbei können die Funktionen der Gene, die für Dehydrogenasen, Synthetasen, Hydratasen-Aldolasen, Thiolasen oder Demethylasen kodieren, mittels der oben genannten gentechnischen Methoden inaktiviert werden, so daß die Erzeugung der betreffenden Enzyme blockiert ist. Vorzugsweise handelt es sich um die Gene, die für Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA-Synthetasen, Enoyl-CoA-Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen kodieren. Ganz besonders bevorzugt sind Gene, die die in der EP-A 0845532 angegebenen Aminosäuresequenzen kodieren und/oder deren Allelvariationen kodierenden Nukleotidsequenzen.

Gegenstand der Erfindung sind demgemäß auch Gen-Strukturen zur Herstellung transformierter Organismen und Mutanten.

Vorzugsweise werden Gen-Strukturen zur Gewinnung der Organismen und Mutanten eingesetzt, bei denen die für Dehydrogenasen, Synthetasen, Hydratasen-Aldolasen, Thiolasen oder Demethylasen kodierenden Nukleotidsequenzen inaktiviert sind. Besonders bevorzugt sind Gen-Strukturen, bei dene die für Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA-Synthetasen, Enoyl-CoA-Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen kodierenden Nukleotidsequenzen inaktiviert sind. Ganz besonders bevorzugt sind Gen-Strukturen, die die in den Figuren 1a bis 1r angegebenen Strukturen mit den in den Figuren 2a bis 2r wiedergegebenen Nukleotidsequenzen und/oder deren Allelvariationen kodierenden Nukleotidsequenzen zen aufweisen. Besonders bevorzugt sind hierbei Nukleotidsequenzen von 1 bis 18.

5

10

15

20

25

30

Die Erfindung schließt auch die Teilsequenzen dieser Gen-Strukturen sowie funktionelle Äquivalente ein. Unter funktionellen Äquivalenten sind solche Derivate der DNA zu verstehen, bei denen einzelne Nukleobasen ausgetauscht worden sind (Wobbelaustausche), ohne die Funktion zu ändern. Auch auf Proteinebene können Aminosäuren ausgetauscht werden, ohne daß eine Veränderung der Funktion die Folge ist.

Den Gen-Strukturen können ein oder mehrere DNA-Sequenzen vor- und/oder nachgeschaltet sein. Durch Klonierung der Gen-Strukturen sind Plasmide bzw. Vektoren erhältlich, die zur Transformation und/oder Transfektion eines Organismus und/oder zur konjugativen Übertragung in einen Organismus geeignet sind.

Gegenstand der Erfindung sind ferner Plasmide und/oder Vektoren zur Herstellung der erfindungsgemäßen transformierten Organismen und Mutanten. Diese enthalten demgemäß die beschriebenen Gen-Strukturen. Die vorliegende Erfindung betrifft

WO 00/26355 PCT/EP99/07952

demgemäß auch Organismen, die die genannten Plasmide und/oder Vektoren enthalten.

Die Art der Plasmide und/oder Vektoren hängt von deren Einsatzzweck ab. Um z. B. die intakten Gene des Eugenol- und/oder Ferulasäure-Katabolismus in Pseudomonaden gegen die durch Omega-Elemente inaktivierten Gene austauschen zu können, benötigt man Vektoren, die einerseits in Pseudomonaden übertragen werden können (konjugativ übertragbare Plasmide), andererseits dort jedoch nicht repliziert werden können und somit in Pseudomonaden instabil sind (sogenannte Suizid-Plasmide). DNA-Abschnitte, die mit Hilfe eines solchen Plasmidsystems in Pseudomonaden übertragen werden, können nur erhalten bleiben, wenn sie durch homologe Rekombination in das Genom der Bakterienzelle integriert werden.

Die beschriebenen Gen-Strukturen, Vektoren und Plasmide können zur Herstellung verschiedener transformierter Organismen oder Mutanten verwendet werden. Mittels der gennanten Gen-Strukturen können intakte Nukleinsäuresequenzen gegen veränderte und/oder inaktivierte Gen-Strukturen ausgetauscht werden. In den durch Transformation oder Transfektion oder Konjugation erhältlichen Zellen erfolgt durch homologe Rekombination ein Austausch des intakten Gens gegen die veränderte und/oder inaktivierte Gen-Struktur, wodurch die resultierenden Zellen nur noch über die veränderte und/oder inaktivierte Gen-Struktur im Genom verfügen. So können erfindungsgemäß vorzugsweise Gene derart verändert und/oder inaktiviert werden, daß die betreffenden Organismen Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure zu erzeugen vermögen.

25

30

5

10

15

20

Erfindungsgemäß derart konstruierte Produktionsstämme sind beispielsweise Mutanten des Stammes *Pseudomonas* sp. HR199 (DSM 7063) der in der DE-A 4 227 076 und der EP-A 0845532 genau beschrieben wurde, wobei sich unter anderem die entsprechenden Genstrukturen aus den Figuren 1a bis 1r in Verbindung mit den Figuren 2a bis 2r ergeben:

- Pseudomonas sp. HR199calAΩKm, enthaltend das durch ΩKm inaktivierte calA-Gen an Stelle des intakten calA-Gens kodierend für Coniferylalkohol-Dehydrogenase (Fig. 1a; Fig. 2a).
- Pseudomonas sp. HR199calAΩGm, enthaltend das durch ΩGm inaktivierte calA-Gen an Stelle des intakten calA-Gens kodierend für Coniferylalkohol-Dehydrogenase (Fig. 1b; Fig. 2b).

15

- 3. Pseudomonas sp. HR199calAΔ, enthaltend das durch Deletion inaktivierte calA-Gen an Stelle des intakten calA-Gens kodierend für Coniferylalkohol-Dehydrogenase (Fig. 1c; Fig. 2c).
- Pseudomonas sp. HR199calBΩKm, enthaltend das durch ΩKm inaktivierte calB-Gen an Stelle des intakten calB-Gens kodierend für Coniferylaldehyd-Dehydrogenase (Fig. 1d; Fig. 2d)
  - 5. Pseudomonas sp. HR199 $calB\Omega$ Gm, enthaltend das durch  $\Omega$ Gm inaktivierte calB-Gen an Stelle des intakten calB-Gens kodierend für Coniferylaldehyd-Dehydrogenase (Fig. 1e; Fig. 2e).
  - 6. Pseudomonas sp. HR199calBΔ, enthaltend das durch Deletion inaktivierte calB-Gen an Stelle des intakten calB-Gens kodierend für Coniferylaldehyd-Dehydrogenase(Fig.1f; Fig. 2f).
- 7. Pseudomonas sp. HR199fcsΩKm, enthaltend das durch ΩKm inaktivierte fcs 20 Gen an Stelle des intakten fcs-Gens kodierend für Ferulasäure-CoA-Synthetase (Fig.1g; Fig. 2g).
  - 8. Pseudomonas sp. HR199 $fcs\Omega$ Gm, enthaltend das durch  $\Omega$ Gm inaktivierte fcs-Gen an Stelle des intakten fcs-Gens kodierend für Ferulasäure-CoA-Synthetase (Fig.1h; Fig. 2h).
- 9. Pseudomonas sp. HR199fcsΔ, enthaltend das durch Deletion inaktivierte fcs-Gen an Stelle des intakten fcs-Gens kodierend für Ferulasäure-CoA-Synthetase (Fig.1i; Fig. 2i).
  - 10. Pseudomonas sp. HR199echΩKm, enthaltend das durch ΩKm inaktivierte ech-Gen an Stelle des intakten ech-Gens kodierend für Enoyl-CoA-Hydratase-Aldolase (Fig.1j; Fig. 2j).

- Pseudomonas sp. HR199echΩGm, enthaltend das durch ΩGm inaktivierte ech-Gen an Stelle des intakten ech-Gens kodierend für Enoyl-CoA-Hydratase-Aldolase (Fig.1k; Fig. 2k).
- 12. Pseudomonas sp. HR199echΔ, enthaltend das durch Deletion inaktivierte ech-Gen an Stelle des intakten ech-Gens kodierend für Enoyl-CoA-Hydratase-Aldolase (Fi.11; Fig. 21).

15

- 13. Pseudomonas sp. HR199aatΩKm, enthaltend das durch ΩKm inaktivierte aat-Gen an Stelle des intakten aat-Gens kodierend für beta-Ketothiolase (Fig. 1m; Fig. 2m).
- 10 14. Pseudomonas sp. HR199aatΩGm, enthaltend das durch ΩGm inaktivierte aat-Gen an Stelle des intakten aat-Gens kodierend für beta-Ketothiolase (Fig.1n; Fig. 2n).
  - Pseudomonas sp. HR199aatΔ, enthaltend das durch Deletion inaktivierte aat-Gen an Stelle des intakten aat-Gens kodierend für beta-Ketothiolase (Fig.1o; 2o).
  - 16. Pseudomonas sp. HR199vdhΩKm, enthaltend das durch ΩKm inaktivierte vdh-Gen an Stelle des intakten vdh-Gens kodierend für Vanillin-Dehydrogenase (Fig.1p; Fig. 2p).
- 17. Pseudomonas sp. HR199vdhΩGm, enthaltend das durch ΩGm inaktivierte
   20 vdh-Gen an Stelle des intakten vdh-Gens kodierend für Vanillin-Dehydrogenase (Fig.1q; Fig. 2q).
  - 18. *Pseudomonas* sp. HR199vdhΔ, enthaltend das durch Deletion inaktivierte vdh-Gen an Stelle des intakten vdh-Gens kodierend für Vanillin-Dehydrogenase (Fig.1r; Fig. 2r).
- 25 19. Pseudomonas sp. HR199vdhBΩKm, enthaltend das durch ΩKm inaktivierte vdhB-Gen an Stelle des intakten vdhB-Gens kodierend für Vanillin-Dehydrogenase II.
  - 20. Pseudomonas sp. HR199vdhBΩGm, enthaltend das durch ΩGm inaktivierte vdhB-Gen an Stelle des intakten vdhB-Gens kodierend für Vanillin-Dehydrogenase II.

- 21. Pseudomonas sp. HR199vdhBΔ, enthaltend das durch Deletion inaktivierte vdhB-Gen an Stelle des intakten vdhB-Gens kodierend für Vanillin-Dehydrogenase II.
- 22. Pseudomonas sp. HR199adhΩKm, enthaltend das durch ΩKm inaktivierte adh-Gen an Stelle des intakten adh-Gens kodierend für Alkohol-Dehydrogenase.
  - 23. Pseudomonas sp. HR199 $adh\Omega$ Gm, enthaltend das durch  $\Omega$ Gm inaktivierte adh-Gen an Stelle des intakten adh-Gens kodierend für Alkohol-Dehydrogenase.
- 24. Pseudomonas sp. HR199adhΔ enthaltend das durch Deletion inaktivierte adh Gen an Stelle des intakten adh-Gens kodierend für Alkohol-Dehydrogenase.
  - 25. Pseudomonas sp. HR199 $vanA\Omega$ Km, enthaltend das durch  $\Omega$ Km inaktivierte vanA-Gen an Stelle des intakten vanA-Gens kodierend für die  $\alpha$ -Untereinheit der Vanillinsäure-Demethylase.
- 15 26. Pseudomonas sp. HR199vanAΩGm, enthaltend das durch ΩGm inaktivierte vanA-Gen an Stelle des intakten vanA-Gens kodierend für die α-Untereinheit der Vanillinsäure-Demethylase.
  - 27. Pseudomonas sp. HR199vanAΔ, enthaltend das durch Deletion inaktivierte vanA-Gen an Stelle des intakten vanA-Gens kodierend für die α-Untereinheit der Vanillinsäure-Demethylase.
  - 28. Pseudomonas sp. HR199vanBΩKm, enthaltend das durch ΩKm inaktivierte vanB-Gen an Stelle des intakten vanB-Gens kodierend für die β-Untereinheit der Vanillinsäure-Demethylase.
- 29. Pseudomonas sp. HR199vanBΩGm, enthaltend das durch ΩGm inaktivierte
   25 vanB-Gen an Stelle des intakten vanB-Gens kodierend für die β-Untereinheit der Vanillinsäure-Demethylase.
  - 30. Pseudomonas sp. HR199vanBΔ, enthaltend das durch Deletion inaktivierte vanB-Gen an Stelle des intakten vanB-Gens kodierend für die β-Untereinheit der Vanillinsäure-Demethylase.

10

15

Gegenstand der Erfindung ist außerdem ein Verfahren zur biotechnischen Herstellung von organischen Verbindungen. Insbesondere können mit diesem Verfahren Alkohole, Aldehyde und organische Säuren hergestellt werden. Vorzugsweise handelt es sich hierbei um Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und Vanillinsäure.

In dem erfindungsgemäßen Verfahren werden die oben beschriebenen Organismen eingesetzt. Zu den ganz besonders bevorzugten Organismen gehören Bakterien, insbesondere die *Pseudomonas*-Arten. Im einzelnen können die oben genannten *Pseudomonas*-Arten vorzugsweise für folgende Verfahren eingesetzt werden:

- 1. Pseudomonas sp. HR199cal $A\Omega$ Km, Pseudomonas sp. HR199cal $A\Omega$ Gm und Pseudomonas sp. HR199cal $A\Delta$  zur Herstellung von Coniferylalkohol aus Eugenol.
- 2. Pseudomonas sp. HR199cal $B\Omega$ Km, Pseudomonas sp. HR199cal $B\Omega$ Gm und Pseudomonas sp. HR199cal $B\Delta$  zur Herstellung von Coniferylaldehyd aus Eugenol oder Coniferylalkohol.
- Pseudomonas sp. HR199fcsΩKm, Pseudomonas sp. HR199fcsΩGm, Pseudomonas sp. HR199fcsΔ, Pseudomonas sp. HR199echΩKm, Pseudomonas sp. HR199echΩGm und Pseudomonas sp. HR199echΔ zur Herstellung von Ferulasäure aus Eugenol oder Coniferylalkohol oder Coniferylaldehyd.
- Pseudomonas sp. HR199vdhΩKm, Pseudomonas sp. HR199vdhΩGm, Pseudomonas sp. HR199vdhΩ GmvdhBΩKm,
   Pseudomonas sp. HR199vdhΩKmvdhBΩGm, Pseudomonas sp. HR199vdhΩvdhBΩGm und Pseudomonas sp. HR199vdhΔvdhBΩKm zur Herstellung von Vanillin aus Eugenol oder Coniferylalkohol oder Coniferylaldehyd oder Ferulasäure.

5. Pseudomonas sp. HR199vanAΩKm, Pseudomonas sp. HR199vanAΩGm, Pseudomonas sp. HR199vanAΔ, Pseudomonas sp. HR199vanBΩKm, Pseudomonas sp. HR199vanBΩGm und Pseudomonas sp. HR199vanBΔ zur Herstellung von Vanillinsäure aus Eugenol oder Coniferylalkohol oder Coniferylaldehyd oder Ferulasäure oder Vanillin.

Bevorzugtes Substrat ist Eugenol. Jedoch kann der Zusatz weiterer Substrate oder sogar der Austausch des Eugenol gegen ein anderes Substrat möglich sein.

10

5

Als Nährmedium für die erfindungsgemäß eingesetzten Organismen kommen synthetische, halbsynthetische oder komplexe Kulturmedien in Betracht. Diese können kohlenstoffhaltige und stickstoffhaltige Verbindungen, anorganische Salze, gegebenenfalls Spurenelemente sowie Vitamine enthalten.

15

Als kohlenstoffhaltige Verbindungen können Kohlenhydrate, Kohlenwasserstoffe oder organische Grundchemikalien in Betracht kommen. Beispiele für vorzugsweise verwendbare Verbindungen sind Zucker, Alkohole bzw. Zuckeralkohole, organische Säuren oder komplexe Gemische.

20

Als Zucker kommt vorzugsweise Glucose in Betracht. Als organische Säuren können vorzugsweise Zitronensäure oder Essigsäure zum Einsatz kommen. Zu den komplexen Gemischen zählen z. B. Malzextrakt, Hefeextrakt, Casein oder Caseinhydrolysat.

25

Als stickstoffhaltige Substrate kommen anorganische Verbindungen in Betracht. Beispiele hierfür sind Nitrate und Ammoniumsalze. Ebenso können organische Stickstoffquellen zum Einsatz kommen. Hierzu zählen Hefecxtrakt, Sojamehl, Casein, Caseinhydrolysat und Maisquellwasser.

Zu den einsetzbaren anorganischen Salzen zählen beispielsweise Sulfate, Nitrate, Chloride, Carbonate und Phosphate. Als Metalle enthalten die genannte Salze vorzugsweise Natrium, Kalium, Magnesium, Mangan, Calcium, Zink und Eisen.

Die Temperatur für die Kultivierung liegt vorzugsweise im Bereich von 5 bis 100°C. Besonders bevorzugt ist der Bereich von 15 bis 60°C, höchst bevorzugt sind 22 bis 37°C.

Der pH-Wert des Mediums beträgt bevorzugt 2 bis 12. Besonders bevorzugt ist der Bereich von 4 bis 8.

10

15

30

Grundsätzlich können für die Durchführung des erfindungsgemäßen Verfahrens alle dem Fachmann bekannten Bioreaktoren eingesetzt werden. Vorzugsweise kommen alle für Submersverfahren geeigneten Vorrichtungen in Betracht. Das heißt, es können erfindungsgemäß Gefäße ohne oder mit mechanischer Mischeinrichtung eingesetzt werden. Zu ersteren zählen z. B. Schüttelapparaturen, Blasensäulen- oder Schlaufenreaktoren. Zu letzteren gehören vorzugsweise alle bekannten Vorrichtungen mit Rührern in beliebiger Gestaltung.

Das erfindungsgemäße Verfahren kann kontinuierlich oder diskontinuierlich durchgeführt werden. Die Dauer der Fermentation bis zum Erreichen einer maximalen Produktmenge hängt von der speziellen Art des eingesetzten Organismus ab. Grundsätzlich liegen die Zeiten der Fermentation jedoch zwischen 2 und 200 Stunden.

Im folgenden wird die Erfindung unter Bezugnahme auf Beispiele näher erläutert:

Von dem Eugenol verwertenden Stamm *Pseudomonas* sp. HR199 (DSM 7063) wurden gezielt Mutanten erzeugt, wobei spezifisch Gene des Eugenol-Katabolismus durch Insertion von Omega-Elementen oder durch Einführen von Deletionen inaktiviert wurden. Als Omega-Elemente dienten DNA-Abschnitte die für Antibiotikaresistenzen gegen Kanamycin (ΩKm) und Gentamycin (ΩGm) codierten. Diese

Resistenzgene wurden ausgehend von Tn5 und dem Plasmid pBBR1MCS-5 mit Hilfe von Standardmethoden isoliert. Die Gene calA, calB, fcs, ech, aat, vdh, adh, vdhB, vanA und vanB, die für Coniferylalkohol-Dehydrogenase, Coniferylaldehyd-Dehydrogenase, Ferulasäure-CoA Synthetase, Enoyl-CoA Hydratase-Aldolase, beta-Ketothiolase, Vanillin-Dehydrogenase, Alkohol-Dehydrogenase, Vanillin-Dehydrogenase II und Vanillinsäure-Demethylase codieren wurden ausgehend von genomischer DNA des Stammes Pseudomonas sp. HR199 mit Hilfe von Standardmethoden isoliert und in pBluescript SK kloniert. Aus diesen Genen wurden durch Verdauung mit geeigneten Restriktionsendonukleasen DNA-Abschnitte entfernt (Deletion), bzw durch  $\Omega$ -Elemente substituiert (Insertion), wodurch das jeweilige Gen inaktiviert wurde. Die auf diese Weise mutierten Gene wurden in konjugativ übertragbare Vektoren umkloniert und anschließend in den Stamm Pseudomonas sp. HR199 eingeführt. Durch geeignete Selektion wurden Transkonjuganten erhalten, die das jeweils funktionsfähige wildtyp-Gen gegen das neu eingebrachte inaktivierte Gen ausgetauscht hatten. Die so erhaltenen Insertions- und Deletionsmutanten wiesen nur noch das jeweils inaktivierte Gen auf. Auf diese Weise wurden sowohl Mutanten mit nur einem defekten Gen als auch Mehrfachmutanten, in denen mehrere Gene auf diese Weise inaktiviert wurden, erhalten. Diese Mutanten wurden für die Biotransformation von

- a) Eugenol zu Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder
   Vanillinsäure;
  - b) Coniferylalkohol zu Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure:
  - c) Coniferylaldehyd zu Ferulasäure, Vanillin und/oder Vanillinsäure;
- d) Ferulasäure zu Vanillin und/oder Vanillinsäure und
  - e) Vanillin zu Vanillinsäure eingesetzt.

5

10

10

15

25

#### Material und Methoden

#### Wachstumsbedingungen der Bakterien.

Stämme von Escherichia coli wurden bei 37°C in Luria-Bertani (LB) oder M9-Mineralmedium (Sambrook, J., E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.) angezogen. Stämme von *Pseudomonas* sp. wurden bei 30°C in Nutrient Broth (NB, 0,8%, wt/vol) oder in Mineralmedium (MM) (Schlegel, H. G. et al. 1961. Arch. Mikrobiol. 38:209-222) bzw. HR-Mineralmedium (HR-MM) (Rabenhorst, J. 1996. Appl. Microbiol. Biotechnol. 46:470-474.) angezogen. Ferulasäure, Vanillin, Vanillinsäure und Protocatechusäure wurden in Dimethylsulfoxid gelöst, und dem jeweiligen Medium in einer Endkonzentration von 0.1% (wt/vol) zugesetzt. Eugenol wurde dem Medium direkt in einer Endkonzentration von 0.1% (vol/vol) zugesetzt, bzw. in den Deckel von MM-Agarplatten auf Filterpapier (Rundfilter 595, Schleicher & Schuell, Dassel, Deutschland) appliziert. Bei der Anzucht von Transkonjuganten und Mutanten von *Pseudomonas* sp. wurde Tetracyclin, Kanamycin, und Gentamycin in Endkonzentrationen von 25 μg/ml bzw. 100 μg/ml bzw. 7,5 μg/ml eingesetzt.

# Qualitativer und quantitativer Nachweis von Stoffwechselintermediaten in Kulturüberständen.

Kulturüberstände wurden direkt, bzw. nach Verdünnung mit H2O-bidest. mittels Hochdruck-Flüssigkeits-Chromatographie (Knauer-HPLC) analysiert. Die Chromatographie erfolgte an Nucleosil-100 C18 (7 μm, 250 x 4 mm). Als Lösungsmittel diente 0.1% (vol/vol) Ameisensäure und Acetonitril. Der verwendete Gradient zur Elution der Substanzen verlief wie folgt:

 $00:00 - 06:30 \rightarrow 26\%$  Acetonitril

 $06:30 - 08:00 \rightarrow 100\%$  Acetonitril

30  $08:00 - 12:00 \rightarrow 100\%$  Acetonitril

12:00 - 13:00 → 26% Acetonitril

13:00 - 18:00 → 26% Acetonitril

## Reinigung der Vanillin-Dehydrogenase-II.

Die Aufreinigung erfolgte bei 4°C.

#### 5 Rohextrakt.

. 10

15

20

Auf Eugenol angezogene Zellen von *Pseudomonas* sp. HR199 wurden in 10 mM Natriumphosphat-Puffer, pH 6.0 gewaschen, im gleichen Puffer resuspendiert und durch zweimalige Passage einer French-Presse (Amicon, Silver Spring, Maryland, USA) bei einem Druck von 1000 psi aufgeschlossen. Das Zellhomogenat wurde einer Ultrazentrifugation (1 h, 100 000 x g, 4°C) unterzogen, wodurch die lösliche Fraktion des Rohextraktes als Überstand erhalten wurde.

## Anionenaustauschchromatographie an DEAE-Sephacel.

Die lösliche Fraktion des Rohextraktes wurde über Nacht gegen 10 mM Natriumphosphat-Puffer, pH 6.0 dialysiert. Das Dialysat wurde auf eine mit 10 mM
Natriumphosphat-Puffer, pH 6.0 äquilibrierte DEAE-Sephacel-Säule (2,6 cm x 35
cm, Bettvolumen[BV]: 186 ml) mit einer Durchflußrate von 0.8 ml/min aufgetragen.
Die Säule wurde mit zwei BV 10 mM Natriumphosphat-Puffer, pH 6.0 gespült. Die
Elution der Vanillin-Dehydrogenase-II (VDH-II) erfolgte mit einem linearen Salzgradient von 0 bis 400 mM NaCl in 10 mM Natriumphosphat-Puffer, pH 6.0 (750
ml). Es wurden 10 ml-Fraktionen aufgefangen. Fraktionen mit hoher VDH-II-Aktivität wurden zum DEAE-Pool vereinigt.

## Bestimmung der Vanillin-Dehydrogenase-Aktivität.

Die Bestimmung der VDH-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test. Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.1 mmol Kalium-Phosphat (pH 7.1), 0.125 μmol Vanillin, 0.5 μmol NAD, 1.2 μmol Pyruvat (Na-Salz), Lactat-Dehydrogenase (1 U; aus Schweineherz) und Enzymlösung. Die Oxidation von Vanillin wurde bei λ = 340 nm verfolgt (ε<sub>Vanillin</sub> = 11,6 cm²/μmol).

Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1 μmol Vanillin pro Minute umsetzt. Die Proteinkonzentrationen in

den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. 193:265-275) bestimmt.

PCT/EP99/07952

#### Bestimmung der Coniferylalkohol-Dehydrogenase-Aktivität.

Die Bestimmung der CADH-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test nach Jaeger et al. (Jaeger, E., L. Eggeling und H. Sahm. 1981. Current Microbiology. 6:333-336). Der Reaktionsansatz mit einem Volumen von I ml enthielt 0.2 mmol Tris/HCl (pH 9.0), 0.4 μmol Coniferylalkohol, 2 μmol NAD, 0.1 mmol Semicarbazid und Enzymlösung. Die Reduktion von NAD wurde bei λ = 340 nm verfolgt (ε = 6.3 cm²/μmol). Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei I U der Enzymmenge entspricht, die I μmol Substrat pro Minute umsetzt. Die Proteinkonzentrationen in den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. 193:265-275) bestimmt.

15

20

25

30

#### Bestimmung der Coniferylaldehyd-Dehydrogenase-Aktivität.

Die Bestimmung der CALDH-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test. Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.1 mmol Tris/HCl (pH 8.8), 0.08  $\mu$ mol Coniferylaldehyd, 2.7  $\mu$ mol NAD und Enzymlösung. Die Oxidation von Coniferylaldehyd zu Ferulasäure wurde bei  $\lambda$  = 400 nm verfolgt ( $\epsilon$  = 34 cm²/ $\mu$ mol). Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1  $\mu$ mol Substrat pro Minute umsetzt. Die Proteinkonzentrationen in den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. 193:265-275) bestimmt.

## Bestimmung der Ferulasäure-CoA-Synthetase (Ferulasäure-Thiokinase)-Aktivität.

Die Bestimmung der FCS-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test, modifiziert nach Zenk et al. (Zenk et al. 1980. Anal. Biochem. 101:182-187). Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.09 mmol Kalium-

Phosphat (pH 7.0), 2.1  $\mu$ mol MgCl<sub>2</sub>, 0.7  $\mu$ mol Ferulasäure, 2  $\mu$ mol ATP, 0.4  $\mu$ mol Coenzym A und Enzymlösung. Die Entstehung des CoA-Esters aus Ferulasäure wurde bei  $\lambda = 345$  nm verfolgt ( $\epsilon = 10$  cm<sup>2</sup>/ $\mu$ mol). Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1  $\mu$ mol Substrat pro Minute umsetzt. Die Proteinkonzentrationen in den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. 193:265-275) bestimmt.

#### Electrophoretische Methoden.

5

10

15

20

25

30

Die Auftrennung von proteinhaltigen Extrakten erfolgte in 7.4% (wt/vol) Polyacrylamidgelen unter nativen Bedingungen nach der Methode von Stegemann et al. (Stegemann et al. 1973. Z. Naturforsch. 28c:722-732) und unter denaturierenden Bedingungen in 11.5% (wt/vol) Polyacrylamidgelen nach der Methode von Laemmli (Laemmli, U. K. 1970. Nature (London) 227:680-685). Zur unspezifischen Proteinfärbung wurde Serva Blue R verwendet. Zur spezifischen Anfärbung der Coniferylalkohol-, Coniferylaldehyd- und Vanillin-Dehydrogenase wurden die Gele für 20 min in 100 mM KP-Puffer (pH 7.0) umgepuffert und anschließend bei 30°C im gleichen Puffer dem 0.08% (wt/vol) NAD, 0.04% (wt/vol) p-Nitroblau-Tetrazolium-chlorid, 0.003% (wt/vol) Phenazine-Methosulfat und 1 mM des jeweiligen Substrates zugesetzt worden war inkubiert, bis entsprechende Farbbanden sichtbar wurden.

Transfer von Proteinen aus Polyacrylamidgelen auf PVDF-Membranen. Proteine wurden aus SDS-Polyacrylamidgelen mit Hilfe eines Semidry-Fastblot Gerätes (B32/33, Biometra, Göttingen, Deutschland) nach Herstellerangaben auf PVDF-Membranen (Waters-Milipore, Bedford, Mass., USA) übertragen.

## Bestimmung von N-terminalen Aminosäuresequenzen.

Die Bestimmung von N-terminalen Aminosäuresequenzen erfolgte mit Hilfe eines Protein Peptide Sequenzers (Typ 477 A, Applied Biosystems, Foster City, USA) und eines PTH-Analysers nach Herstellerangaben.

## Isolierung und Manipulation von DNA.

Die Isolierung von genomischer DNA erfolgte nach der Methode von Marmur (Marmur, J. 1961. J. Mol. Biol. 3:208-218). Die Isolierung und Analyse von anderer Plasmid-DNA bzw. von DNA-Restriktionsfragmenten erfolgte nach Standardmethoden (Sambrook, J. E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Habor, New York.).

#### Transfer von DNA.

5

Die Präparation und Transformation von kompetenten Escherichia coli-Zellen 10 erfolgte nach der Methode von Hanahan (Hanahan, D. 1983. J. Mol. Biol. 166:557-580). Konjugativer Plasmidtransfer zwischen Plasmid-tragenden Escherichia coli S17-1-Stämmen (Donor) und Pseudomonas sp.-Stämmen (Rezipient) erfolgte auf NB-Agarplatten nach der Methode von Friedrich et al. (Friedrich, B. et al. 1981. J. Bacteriol. 147:198-205), oder durch eine "Minikomplementations-Methode" auf 15 MM-Agarplatten mit 0.5% (wt/vol) Gluconat als C-Quelle und 25 µg/ml Tetracyclin oder 100 µg/ml Kanamycin. Dabei wurden Zellen des Rezipienten in einer Richtung als Impfstrich aufgetragen. Nach 5 min wurden dann Zellen der Donor-Stämme als Impfstriche aufgetragen, wobei der Rezipienten-Impfstrich gekreuzt wurde. Nach einer Inkubation für 48 h bei 30°C wuchsen die Transkonjuganten direkt hinter der 20 Kreuzungsstelle, wohingegen weder Donor- noch Rezipienten-Stamm zum Wachstum in der Lage war.

#### Hybridisierungsexperimente.

DNA-Restriktionsfragmente wurden in einem 0.8% (wt/vol) Agarose-Gel in 50 mM

Tris- 50 mM Borsäure- 1.25 mM EDTA-Puffer (pH 8.5) elektrophoretisch aufgetrennt (Sambrook, J. E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Habor, New York.) Die Übertragung der denaturierten DNA aus dem Gel auf eine positiv geladene Nylonmembran (Porengröße: 0.45 μm, Pall Filtrationstechnik, Dreieich, Deutschland), die anschließende Hybridisierung mit biotimylierten, bzw.

Digoxigenin-markierten DNA-Sonden und die Herstellung dieser DNA-Sonden erfolgte nach Standardmethoden (Sambrook, J. E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Habor, New York.).

5

10

#### DNA-Sequenzierung.

Die Bestimmung von Nukleotidsequenzen erfolgte nach der Didesoxy-Kettenabbruch-Methode von Sanger et al. (Sanger et al. 1977. Proc. Natl. Acad. Sci. USA 74:5463-5467) "nicht-radioaktiv" mit einem "LI-COR DNA-Sequencer Modell 4000L" (LI-COR Inc., Biotechnology Division, Lincoln, NE, USA) unter Verwendung eines "Thermo Sequenase fluorescent labelled primer cycle sequencing kit with 7-deaza-dGTP" (Amersham Life Science, Amersham International pls, Little Chalfont, Buckinghamshire, England) jeweils nach Vorschrift des Herstellers.

Mit Hilfe von synthetischen Oligonukleotiden wurde nach der "Primer-hopping Strategie" von Strauss et al. (Strauss, E. C. et al. 1986. Anal. Biochem. 154:353-360) sequenziert.

## Chemikalien, Biochemikalien und Enzyme.

Restriktionsenzyme, T4 DNA-Ligase, Lambda-DNA und Enzyme bzw. Substrate für die optisch enzymatischen Tests wurden von C. F. Boehringer & Söhne (Mannheim, Deutschland) oder von GIBCO/BRL (Eggenstein, Deutschland) bezogen. [γ-32P]ATP kam von Amersham/Buchler (Braunschweig, Deutschland). Oligonukleotide wurden von der Firma MWG-Biotech GmbH (Ebersberg, Deutschland) bezogen. Agarose vom Typ NA wurde von Pharmacia-LKB (Uppsala, Schweden) bezogen. Alle anderen Chemikalien waren von Haarmann & Reimer (Holzminden, Deutschland), E. Merck AG (Darmstadt, Deutschland), Fluka Chemie (Buchs,, Schweiz), Serva Feinbiochemica (Heidelberg, Deutschland) oder Sigma Chemie (Deisenhofen, Deutschland).

#### Beispiele

#### Beispiel 1

10

15

Konstruktion von Omega-Elementen, die Resistenzen gegenüber Kanamycin ( $\Omega$  Km) bzw. Gentamycin( $\Omega$ Gm) vermitteln.

Für die Konstruktion des ΩKm-Elements wurde das 2099 bp *BgI*I-Fragment des Transposons Tn5 (Auerswald E. A., G. Ludwig und H. Schaller. 1981. Cold Spring Harb. Symp. Quant. Biol. 45:107-113; Beck E., G. Ludwig, E. A. Auerswald, B. Reiss und H. Schaller. 1982. Gene 19:327-336; Mazodier P., P. Cossart, E. Giraud und F. Gasser. 1985. Nucleic Acids Res. 13:195-205.) präparativ isoliert. Das Fragment wurde durch Behandlung mit der Nuklease Bal-31 auf ca. 990 bp verkürzt. Dieses Fragment, das nur noch das Kanamycin-Resistenzgen (codierend für eine Aminoglycosid-3'-O-Phosphotransferase) umfaßte, wurde anschließend mit *Sma*I geschnittener pSKsym-DNA (pBluescript SK<sup>-</sup>-Derivat, welches eine symetrisch aufgebaute multiple Klonierungsstelle [*Sal*I, *Hin*dIII, *Eco*RI, *Sma*I, *Eco*RI, *Hin*dIII, *Sal*I] enthält) ligiert. Aus dem resultierenden Plasmid konnte das ΩKm-Element als *Sma*I-, *Eco*RI-, *Hin*dIII- oder *Sal*I-Fragment reisoliert werden.

Für die Konstruktion des ΩGm-Elements wurde das 983 bp Eael-Fragment des Plasmids pBBR1MCS-5 (Kovach M.E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop und K. M. Peterson. 1995. Gene 166:175-176.) präparativ isoliert und anschließend mit Mung Bean Nuklease (Abdauen von einzelsträngigen DNA-Molekülenden) behandelt. Dieses Fragment, das nur noch das Gentamycin-Resistenzgen (codierend für eine Gentamycin-3-Acetyltransferase) umfaßte, wurde anschließend mit Smal geschnittener pSKsym-DNA (s.o.) ligiert. Aus dem resultierenden Plasmid konnte das ΩGm-Element als Smal-, EcoRI-, HindIII- oder Sall-Fragment reisoliert werden.

## Beispiel 2

Klonierung der Gene aus *Pseudomonas* sp. HR199 (DSM7063), die durch Insertion von  $\Omega$ -Elementen oder durch Deletionen inaktiviert werden sollten.

Die separaten Klonierungen der Gene fcs, ech, vdh und aat erfolgte ausgehend von den E. coli S17-1 Stämmen DSM 10439 und DSM 10440 mit den Plasmiden pE207 und pE5-1 (siehe EP-A 0845532). Aus diesen Plasmiden wurden die angegebenen Fragmente präparativ isoliert und wie im weiteren beschrieben behandelt:

Für die Klonierung des fcs-Gens wurde das 2350 bp große Sall/EcoRI-Fragment des Plasmids pE207 und das 3700 bp große EcoRI/SalI-Fragment des Plasmids pE5-1 zusammen in pBluescript SK in einer Weise kloniert, daß beide Fragmente über die EcoRI-Enden miteinander verbunden waren. Ausgehend von dem resultierenden Hybridplasmid wurde das 6050 bp SalI-Fragment präparativ isoliert und durch Behandlung mit der Nuklease Bal-31 auf ca. 2480 bp verkürzt. Anschließend wurden an die Fragment-Enden PstI-Linker ligiert und das Fragment nach PstI-Verdauung in pBluescript SK kloniert (pSKfcs). Nach Transformation von E. coli XL1-Blue wurden Klone erhalten, die das fcs-Gen exprimierten und eine FCS-Aktivität von 0.2 U/mg Protein aufwiesen.

20

25

30

5

Für die Klonierung des *ech*-Gens wurde das 3800 bp große *HindIII/Eco*RI-Fragment des Plasmids pE207 präparativ isoliert und durch Behandlung mit der Nuklease Bal-31 auf ca. 1470 bp verkürzt. Anschließend wurden an die Fragment-Enden *Eco*RI-Linker ligiert und das Fragment nach *Eco*RI-Verdauung in pBluescript SK<sup>-</sup> kloniert (pSK*ech*).

Für die Klonierung des *vdh*-Gens wurde das 2350 bp große *Sall/Eco*RI-Fragment des Plasmids pE207 präparativ isoliert. Nach Klonierung in pBluescript SK<sup>-</sup> wurde das Fragment mit Hilfe eines Exonuklease III/ Mung Bean Nukleasesystems einseitig um ca. 1530 bp verkürzt. Anschließend wurde an das Fragmentende ein *Eco*RI-

Linker ligiert und das Fragment nach EcoRI-Verdauung in pBluescript SK kloniert (pSKvdh). Nach Transformation von E. coli XL1-Blue wurden Klone erhalten, die das vdh-Gen exprimierten und eine VDH-Aktivität von 0.01 U/mg Protein aufwiesen.

5

10

15

20

25

30

Für die Klonierung des *aat*-Gens wurde das 3700 bp große *EcoRI/SaI*I-Fragment des Plasmids pE5-1 präparativ isoliert und durch Behandlung mit der Nuklease Bal-31 auf ca. 1590 bp verkürzt. Anschließend wurden an die Fragment-Enden *EcoRI*-Linker ligiert und das Fragment nach *EcoRI*-Verdauung in pBluescript SK kloniert (pSK*aat*).

#### Beispiel 3

Inaktivierung der oben beschriebenen Gene durch Insertion von  $\Omega$ -Elementen, bzw durch Deletion von Teilbereichen dieser Gene.

Das Plasmid pSKfcs, welches das fcs-Gen enthielt wurde mit BssHII verdaut, wodurch ein 1290 bp großes Fragment aus dem fcs-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des fcs-Gens ( $fcs\Delta$ ) (siehe Abb. 1i und 2i) kloniert in pBluescript SK $^-$  (pSK $fcs\Delta$ ) erhalten. Darüber hinaus wurden nach Herausschneiden des Fragments die Omega-Elemente  $\Omega$ Km und  $\Omega$ Gm an dessen Stelle einligiert. Dadurch entstanden die  $\Omega$ -inaktivierten Derivate des fcs-Gens ( $fcs\Omega$ Km, siehe Abb. 1g und 2g) und ( $fcs\Omega$ Gm, siehe Abb. 1h und 2h) kloniert in pBluescript SK $^-$  (pSK $fcs\Omega$ Km und pSK $fcs\Omega$ Gm). In Rohextrakten der erhaltenen E. coli Klone, deren Hybridplasmide ein durch Deletion bzw.  $\Omega$ -Element-Insertion inaktiviertes fcs-Gen aufwiesen, konnte keine FCS-Aktivität nachgewiesen werden.

Das Plasmid pSKech, welches das ech-Gen enthielt, wurde mit Nrul verdaut, wodurch ein 53 bp und ein 430 bp großes Fragment aus dem ech-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des ech-Gens (ech $\Delta$ , siehe Abb. 11 und 21) kloniert in pBluescript SK (pSKech $\Delta$ ) erhalten. Darüber

hinaus wurden nach Herausschneiden der Fragmente die Omega-Elemente  $\Omega$ Km und  $\Omega$ Gm an deren Stelle einligiert. Dadurch entstanden die  $\Omega$ -inaktivierten Derivate des ech-Gens (ech $\Omega$ Km und ech $\Omega$ Gm) kloniert in pBluescript SK $^-$  (pSKech $\Omega$ Km und pSKech $\Omega$ Gm).

5

10

15

Das Plasmid pSKvdh, welches das vdh-Gen enthielt wurde mit BssHII verdaut, wodurch ein 210 bp großes Fragment aus dem vdh-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des vdh-Gens ( $vdh\Delta$ , siehe Abb. 10 und 20) kloniert in pBluescript SK $^-$  (pSK $vdh\Delta$ ) erhalten. Darüber hinaus wurden nach Herausschneiden des Fragments die Omega-Elemente  $\Omega$ Km und  $\Omega$ Gm an dessen Stelle einligiert. Dadurch entstanden die  $\Omega$ -inaktivierten Derivate des vdh-Gens ( $vdh\Omega$ Km und  $vdh\Omega$ Gm) kloniert in pBluescript SK $^-$  (pSK $vdh\Omega$ Km, siehe Abb. 1m und 2m) und (pSK $vdh\Omega$ Gm, siehe Abb. 1n und 2n). In Rohextrakten der erhaltenen E. coli Klone, deren Hybridplasmide ein durch Deletion bzw.  $\Omega$ -Element-Insertion inaktiviertes vdh-Gen aufwiesen, konnte keine VDH-Aktivität nachgewiesen werden.

20

25

Das Plasmid pSKaat, welches das aat-Gen enthielt wurde mit BssHII verdaut, wodurch ein 59 bp großes Fragment aus dem aat-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des aat-Gens (aatΔ, siehe Abb. 1r und 2r) kloniert in pBluescript SK (pSKaatΔ) erhalten. Darüber hinaus wurden nach Herausschneiden des Fragments die Omega-Elemente  $\Omega$ Km und  $\Omega$ Gm an dessen Stelle einligiert. Dadurch entstanden die  $\Omega$ -inaktivierten Derivate des aat-Gens (aat  $\Omega$ Km, siehe Abb. 1p und 2p) und (aat $\Omega$ Gm, siehe Abb. 1q und 2q) kloniert in pBluescript SK (pSKaat $\Omega$ Km und pSKaat $\Omega$ Gm).

## Beispiel 4

5

10

15

20

Umklonieren der durch  $\Omega$ -Elemente inaktivierten Gene in das konjugativ übertragbare "Suizid-Plasmid" pSUP202.

Um die durch Ω-Elemente inaktivierten Gene in *Pseudomonas* sp. HR199 gegen die intakten Gene austauschen zu können, benötigt man einen Vektor, der einerseits in Pseudomonaden übertragen werden kann (konjugativ übertragbare Plasmide), andererseits dort jedoch nicht repliziert werden kann und somit in Pseudomonaden instabil ist ("Suizid-Plasmid"). DNA-Abschnitte, die mit Hilfe eines solchen Plasmidsystems in Pseudomonaden übertragen werden, können nur erhalten bleiben, wenn sie durch homologe Rekombination (RecA-abhängige Rekombination) in das Genom der Bakterienzelle integriert werden. Im vorliegenden Fall wurde das "Suizid-Plasmid" pSUP202 (Simon et al. 1983. *In*: A. Pühler. Molecular genetics of the bacteria-plant interaction. Springer Verlag, Berlin, Heidelberg, New York, S. 98-106.) eingesetzt.

Die inaktivierten Gene  $fcs\Omega Km$  und  $fcs\Omega Gm$  wurden nach PstI-Verdauung aus den Plasmiden pSK $fcs\Omega Km$  und pSK $fcs\Omega Gm$  isoliert und mit PstI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren Hybridplasmid (pSUP $fcs\Omega Km$ ) das inaktivierte Gen  $fcs\Omega Km$  enthielt. Das entsprechende Hybridplasmid (pSUP $fcs\Omega Gm$ ) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen  $fcs\Omega Gm$ .

25

30

Die inaktivierten Gene  $ech\Omega$ Km und  $ech\Omega$ Gm wurden nach EcoRI-Verdauung aus den Plasmiden pSK $ech\Omega$ Km und pSK $ech\Omega$ Gm isoliert und mit EcoRI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren

Hybridplasmid (pSUPech $\Omega$ Km) das inaktivierte Gen ech $\Omega$ Km enthielt. Das entsprechende Hybridplasmid (pSUPech $\Omega$ Gm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen ech $\Omega$ Gm.

Die inaktivierten Gene vdhΩKm und vdhΩGm wurden nach EcoRI-Verdauung aus den Plasmiden pSKvdhΩKm und pSKvdhΩGm isoliert und mit EcoRI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren Hybridplasmid (pSUPvdhΩKm) das inaktivierte Gen vdhΩKm enthielt. Das entsprechende Hybridplasmid (pSUPvdhΩGm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen vdhΩGm.

Die inaktivierten Gene  $aat\Omega$ Km und  $aat\Omega$ Gm wurden nach EcoRI-Verdauung aus den Plasmiden pSK $aat\Omega$ Km und pSK $aat\Omega$ Gm isoliert und mit EcoRI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren Hybridplasmid (pSUP $aat\Omega$ Km) das inaktivierte Gen  $aat\Omega$ Km enthielt. Das entsprechende Hybridplasmid (pSUP $aat\Omega$ Gm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen  $aat\Omega$ Gm.

## Beispiel 5

15

20

30

Umklonieren der durch Deletion inaktivierten Gene in das konjugativ übertragbare "Suizid-Plasmid" mit "sacB-Selektionssystem" pHE55.

Um die durch Deletion inaktivierten Gene in *Pseudomonas* sp. HR199 gegen die intakten Gene austauschen zu können, benötigt man einen Vektor, der die schon für pSUP202 beschriebenen Eigenschaften aufweist. Da im Gegensatz zu den durch  $\Omega$ -Element inaktivierten Genen bei durch Deletion inaktivierten Genen keine Selek-

10

15

20

25

30

tionsmöglichkeit (keine Antibiotika-Resistenz) für den erfolgten Austausch der Gene in Pseudomonas sp. HR199 besteht, mußte ein anderes Selektionssystem zur Anwendung kommen. Bei dem "sacB-Selektionssystem" wird das auszutauschende. durch Deletion inaktivierte Gen in einem Plasmid kloniert, welches neben einem Antibiotika-Resistenzgen auch über das sacB-Gen verfügt. Nach konjugativer Übertragung dieses Hybridplasmids in einen Pseudomonaden wird das Plasmid durch homologe Rekombination an der Stelle in das Genom integriert, an der sich das intakte Gen befindet (erster "Cross over"). Auf diese Weise entsteht ein "heterogenoter" Stamm, der sowohl über ein intaktes als auch über ein durch Deletion inaktiviertes Gen verfügt, welche durch die pHE55-DNA voneinander getrennt sind. Diese Stämme weisen die durch den Vektor codierte Resistenz auf und besitzen darüber hinaus ein aktives sacB-Gen. Durch ein zweites homologes Rekombinationsereignis (zweiter "Cross over"), soll nun die pHE55-DNA zusammen mit dem intakten Gen aus der genomischen DNA ausgegliedert werden. Durch dieses Rekombinationsereignis entsteht ein Stamm, der nur noch über das inaktivierte Gen verfügt. Darüber hinaus kommt es zum Verlust der pHE55-codierten Antibiotika-Resistenz und des sacB-Gens. Streicht man Stämme auf Saccharose-haltigen Medien aus, werden Stämme die das sacB-Gen exprimieren im Wachstum gehemmt, da das Genprodukt Saccharose zu einem Polymer umsetzt, welches im Periplasma der Zellen akkumuliert wird. Zellen, die durch das zweite Rekombinationsereignis das sacB-Gen nicht mehr tragen, werden somit nicht im Wachstum gehemmt. Um eine phänotypische Selektionsmöglichkeit auf die Integration des durch Deletion inaktivierten Gens zu besitzen, tauscht man dieses nicht gegen ein intaktes Gen aus, sondern man bedient sich eines Stammes, in dem das auszutauschende Gen bereits durch Insertion eines Ω-Elements "markiert" vorliegt. Bei erfolgreichem Austausch verliert der resultierende Stamm die durch das Ω-Element codierte Antibiotika-Resistenz.

Das inaktivierte Gen fcsΔ wurden nach Pstl-Verdauung aus dem Plasmid pSKfcsΔ isoliert und mit Pstl geschnittener pHE55 DNA ligiert. Der Ligationsansatz wurde nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclifi-haltigem LB-

Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE $fcs\Delta$ ) das inaktivierte Gen  $fcs\Delta$  enthielt.

Das inaktivierte Gen  $ech\Delta$  wurden nach EcoRI-Verdauung aus dem Plasmid pSK $ech\Delta$  isoliert und mit Mung Bean Nuklease behandelt (Erzeugung von glatten Enden ["blunt ends"]). Das Fragment wurde mit BamHI geschnittener und Mung Bean Nuklease behandelter pHE55 DNA ligiert. Der Ligationsansatz wurde nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE $ech\Delta$ ) das inaktivierte Gen  $ech\Delta$  enthielt.

5

10

15

20

25

Das inaktivierte Gen  $vdh\Delta$  wurden nach EcoRI-Verdauung aus dem Plasmid pSK $vdh\Delta$  isoliert und mit Mung Bean Nuklease behandelt. Das Fragment wurde mit BamHI geschnittener und Mung Bean Nuklease behandelter pHE55 DNA ligiert. Der Ligationsansatz wurde nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE $vdh\Delta$ ) das inaktivierte Gen  $vdh\Delta$  enthielt.

Das inaktivierte Gen  $aat\Delta$  wurden nach EcoRI-Verdauung aus dem Plasmid pSK $aat\Delta$  isoliert und mit Mung Bean Nuklease behandelt. Das Fragment wurde mit BamHI geschnittener und Mung Bean Nuklease behandelter pHE55 DNA ligiert. Der Ligationsansatz wurde nach  $E.\ coli\ S17-1$  transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE $aat\Delta$ )das inaktivierte Gen  $aat\Delta$  enthielt.

## Beispiel 6

Erzeugung von Mutanten des Stammes *Pseudomonas* sp. HR199, bei denen spezifisch Gene des Eugenol-Katabolismuses durch Insertion eines  $\Omega$ -Elementes inaktiviert wurden.

Der Stamm Pseudomonas sp. HR199 wurde als Rezipient in Konjugationsexperimenten eingesetzt, bei denen Stämme von E. coli S17-1 als Donoren eingesetzt wurden, die die unten aufgeführten Hybridplasmide von pSUP202 enthielten. Die Transkonjuganten wurden auf Gluconat-haltigem Mineralmedium selektiert, welches das dem  $\Omega$ -Element entsprechende Antibiotikum enthielt. "Homogenote" (Austausch des intakten Gens gegen das durch  $\Omega$ -Element-Insertion inaktivierte Gen durch doppeltes "Cross over") und "heterogenote" (Integration des Hybridplasmids in das Genom durch einfachen "Cross over") Transkonjuganten konnten anhand der durch pSUP202 codierten Tetracyclin-Resistenz unterschieden werden.

15

5

10

Die Mutanten Pseudomonas sp. HR199  $fcs\Omega$ Km und Pseudomonas sp. HR199  $fcs\Omega$ Gm wurden nach Konjugation von Pseudomonas sp. HR199 mit E. coli S17-1 (pSUP $fcs\Omega$ Km) bzw. E. coli S17-1 (pSUP $fcs\Omega$ Gm) erhalten. Der Austausch des intakten fcs-Gens gegen das durch  $\Omega$ Km bzw.  $\Omega$ Gm inaktivierte Gen ( $fcs\Omega$ Km bzw.  $fcs\Omega$ Gm) wurde mittels DNA-Sequenzierung verifiziert.

20

Die Mutanten Pseudomonas sp. HR199  $ech\Omega$ Km und Pseudomonas sp. HR199  $ech\Omega$ Gm wurden nach Konjugation von Pseudomonas sp. HR199 mit E. coli S17-1 (pSUP $ech\Omega$ Km) bzw. E. coli S17-1 (pSUP $ech\Omega$ Gm) erhalten. Der Austausch des intakten ech-Gens gegen das durch  $\Omega$ Km bzw.  $\Omega$ Gm inaktivierte Gen ( $ech\Omega$ Km bzw.  $ech\Omega$ Gm) wurde mittels DNA-Sequenzierung verifiziert.

25

30

Die Mutanten Pseudomonas sp. HR199  $vdh\Omega$ Km und Pseudomonas sp. HR199  $vdh\Omega$ Gm wurden nach Konjugation von Pseudomonas sp. HR199 mit E. coli S17-1 (pSUP $vdh\Omega$ Km) bzw. E. coli S17-1 (pSUP $vdh\Omega$ Gm) erhalten. Der Austausch des

intakten vdh-Gens gegen das durch  $\Omega$ Km bzw.  $\Omega$ Gm inaktivierte Gen ( $vdh\Omega$ Km bzw.  $vdh\Omega$ Gm) wurde mittels DNA-Sequenzierung verifiziert.

Die Mutanten Pseudomonas sp. HR199  $aat\Omega$ Km und Pseudomonas sp. HR199  $aat\Omega$  Gm wurden nach Konjugation von Pseudomonas sp. HR199 mit E. coli S17-1 (pSUP $aat\Omega$ Km) bzw. E. coli S17-1 (pSUP $aat\Omega$ Gm) erhalten. Der Austausch des intakten aat-Gens gegen das durch  $\Omega$ Km bzw.  $\Omega$ Gm inaktivierte Gen ( $aat\Omega$ Km bzw.  $aat\Omega$ Gm) wurde mittels DNA-Sequenzierung verifiziert.

- Die Mutante *Pseudomonas* sp. HR199 fcsΩKmvdhΩGm wurden nach Konjugation von *Pseudomonas* sp. HR199 fcsΩKm mit E. coli S17-1 (pSUPvdhΩGm) erhalten. Der Austausch des intakten vdh-Gens gegen das durch ΩGm inaktivierte Gen (vdhΩGm) wurde mittels DNA-Sequenzierung verifiziert.
- Die Mutante Pseudomonas sp. HR199 vdhΩKmaatΩGm wurden nach Konjugation von Pseudomonas sp. HR199 vdhΩKm mit E. coli S17-1 (pSUPaatΩGm) erhalten. Der Austausch des intakten aat-Gens gegen das durch ΩGm inaktivierte Gen (aatΩ Gm) wurde mittels DNA-Sequenzierung verifiziert.
- Die Mutante Pseudomonas sp. HR199 vdhΩKmechΩGm wurden nach Konjugation von Pseudomonas sp. HR199 vdhΩKm mit E. coli S17-1 (pSUPechΩGm) erhalten. Der Austausch des intakten ech-Gens gegen das durch ΩGm inaktivierte Gen (echΩGm) wurde mittels DNA-Sequenzierung verifiziert.

#### Beispiel 7

5

10

15

Erzeugung von Mutanten des Stammes *Pseudomonas* sp. HR199, bei denen spezifisch Gene des Eugenol-Katabolismuses durch Deletion eines Teilbereiches inaktiviert wurden.

Die Stämme Pseudomonas sp. HR199  $fcs\Omega$ Km, Pseudomonas sp. HR199  $ech\Omega$ Km, Pseudomonas sp. HR199  $vdh\Omega$ Km und Pseudomonas sp. HR199  $aat\Omega$ Km wurden als Rezipient in Konjugationsexperimenten eingesetzt, bei denen Stämme von E. coli S17-1 als Donoren eingesetzt wurden, die die unten aufgeführten Hybridplasmide von pHE55 enthielten. Die "heterogenoten" Transkonjuganten wurden auf Gluconathaltigem Mineralmedium selektiert, welches neben Tetracyclin (pHE55 codierte Resistenz) das dem  $\Omega$ -Element entsprechende Antibiotikum enthielt. Nach Ausstreichen auf Saccharose-haltigem Mineralmedium wurden Transkonjuganten erhalten, die durch ein zweites Rekombinationsereignis (zweiter "Cross over") die Vektor-DNA eliminiert hatten. Durch Ausstreichen auf Mineralmedium ohne Antibiotika bzw. mit dem  $\Omega$ -Element entsprechenden Antibiotikum konnten die Mutanten erkannt werden, bei denen das durch  $\Omega$ -Element inaktivierte Gen gegen das durch Deletion inaktivierte Gen ausgetauscht worden war (keine Antibiotika-Resistenz).

- Die Mutante Pseudomonas sp. HR199 fcsΔ wurde nach Konjugation von Pseudomonas sp. HR199 fcsΩKm mit E. coli S17-1 (pHEfcsΔ) erhalten. Der Austausch des durch ΩKm inaktivierten Gens (fcsΩKm) gegen das durch Deletion inaktivierte Gen (fcsΔ) wurde mittels DNA-Sequenzierung verifiziert.
- Die Mutanten Pseudomonas sp. HR199 echΔ wurde nach Konjugation von Pseudomonas sp. HR199 echΩKm mit E. coli S17-1 (pHEechΔ) erhalten. Der Austausch des durch ΩKm inaktivierten Gens (echΩKm) gegen das durch Deletion inaktivierte Gen (echΔ) wurde mittels DNA-Sequenzierung verifiziert.

Die Mutanten *Pseudomonas* sp. HR199  $vdh\Delta$  wurde nach Konjugation von *Pseudomonas* sp. HR199  $vdh\Omega$ Km mit *E. coli* S17-1 (pHE $vdh\Delta$ ) erhalten. Der Austausch des durch  $\Omega$ Km inaktivierten Gens ( $vdh\Omega$ Km) gegen das durch Deletion inaktivierte Gen ( $vdh\Delta$ ) wurde mittels DNA-Sequenzierung verifiziert.

5

Die Mutanten *Pseudomonas* sp. HR199  $aat\Delta$  wurde nach Konjugation von *Pseudomonas* sp. HR199  $aat\Omega$ Km mit *E. coli* S17-1 (pHE $aat\Delta$ ) erhalten. Der Austausch des durch  $\Omega$ Km inaktivierten Gens ( $aat\Omega$ Km) gegen das durch Deletion inaktivierte Gen ( $aat\Delta$ ) wurde mittels DNA-Sequenzierung verifiziert.

10

15

25

#### **Beispiel 8**

Biotransformation von Eugenol zu Vanillin mit der Mutante Pseudomonas sp.  $HR199 \ vdh\Omega Km$ .

Der Stamm *Pseudomonas* sp. HR199 *vdh*ΩKm wurde in 50 ml HR-MM mit 6 mM Eugenol bis zu einer optischen Dichte von ca. OD600nm = 0.6 angezogen. Nach 17 h waren 2.9 mM Vanillin, 1.4 mM Ferulasäure und 0.4 mM Vanillinsäure im Kultur-überstand nachweisbar.

#### 20 Beispiel 9

Biotransformation von Eugenol zu Ferulasäure mit der Mutante Pseudomonas sp. HR199  $vdh\Omega Gmaat\Omega Km$ .

Der Stamm *Pseudomonas* sp. HR199 *vdh*ΩGmaatΩKm wurde in 50 ml HR-MM mit 6 mM Eugenol bis zu einer optischen Dichte von ca. OD600nm = 0.6 angezogen. Nach 18 h waren 1.9 mM Vanillin, 2.4 mM Ferulasäure und 0.6 mM Vanillinsäure im Kulturüberstand nachweisbar.

WO 00/26355 PCT/EP99/07952

### Beispiel 10

Biotransformation von Eugenol zu Coniferylalkohol mit der Mutante Pseudomonas sp. HR199  $vdh\Omega Gmaat\Omega Km$ .

Der Stamm *Pseudomonas* sp. HR199 *vdh*ΩGmaatΩKm wurde in 50 ml HR-MM mit 6 mM Eugenol bis zu einer optischen Dichte von ca. OD600nm = 0.4 angezogen. Nach 15 h waren 1.7 mM Coniferylalkohol, 1.4 mM Vanillin, 1.4 mM Ferulasäure und 0.2 mM Vanillinsäure im Kulturüberstand nachweisbar.

#### Beispiel 11

10

15

20

25

Fermentative Produktion von natürlichem Vanillin aus Eugenol im 10 l Fermenter mit der Mutante Pseudomonas sp. HR 199  $vdh\Omega Km$ .

Mit 100 ml einer 24 Stunden alten Vorkultur, die auf einer Schüttelmaschine (120 Upm) bei 32°C in einem auf pH 7,0 eingestellten Medium aus 12,5 g/l Glyzerin, 10 g/l Hefeextrakt und 0,37 g/l Essigsäure angezogen wurde, wurde der Produktionsfermenter beimpft. Der Fermenter enthielt 9,9 l Medium mit folgender Zusammensetzung: 1,5 g/l Hefeextrakt, 1,6 g/l KH<sub>2</sub>PO<sub>4</sub>, 0,2 g/l NaCl, 0,2 g/l MgSO<sub>4</sub>. Der pH-Wert wurde mit Natronlauge auf pH 7,0 eingestellt. Nach der Sterilisation wurde dem Medium 4 g Eugenol zugefügt. Die Temperatur betrug 32°C, die Belüftung 3 Nl/min und die Rührerdrehzahl 600 Upm. Der pH-Wert wurde mit Natronlauge bei pH 6,5 gehalten.

Vier Stunden nach dem Animpfen wurde mit der kontinuierlichen Zugabe von Eugenol begonnen, so daß am Ende der Fermentation nach 65 Stunden 255 g Eugenol zur Kultur gegeben worden waren. Außerdem wurden während der Fermentation 40 g Hefeextrakt zugefüttert. Die Konzentration an Eugenol lag am Ende der Fermentation bei 0,2 g/l. Der Gehalt an Vanillin betrug 2,6 g/l. Zusätzlich lagen noch 3,4 g/l Ferulasäure vor.

Das so erhaltene Vanillin kann durch bekannte physikalische Verfahren wie Chromatographie, Destillation und/oder Extraktion isoliert und zur Herstellung natürlicher Aromen verwendet werden.

# 5 Erläuterungen zu den Figuren:

FIG. 1a bis 1r:

Gen- Strukturen zur Gewinnung von Organismen und Mutanten

10

15

20

calA\*: Teil des inaktivierten Gens der Coniferylalkohol-Dehydrogenase

calB\*: Teil des inaktivierten Gens der Coniferylaldehyd-Dehydrogenase

fcs\*: Teil des inaktivierten Gens der Ferulasäure-CoA Synthetase

ech\*: Teil des inaktivierten Gens der Enoyl-CoA Hydratase-Aldolase

vdh\*: Teil des inaktivierten Gens der Vanillin-Dehydrogenase

aat\*: Teil des inaktivierten Gens der beta-Ketothiolase

Die mit "\*" versehenen Restriktionsenzym-Schnittstellen kamen für die Konstruktion zum Einsatz, sind jedoch in dem resultierenden Konstrukt nicht mehr funktionsfähig.

	FIG. 2a: Nukleotidsequenz der Gen-Struktur calAΩKm
	FIG. 2b: Nukleotidsequenz der Gen-Struktur calAΩGm
	FIG. 2c: Nukleotidsequenz der Gen-Struktur calAΔ
	FIG. 2d: Nukleotidsequenz der Gen-Struktur $calB\Omega$ Km
5	FIG. 2e: Nukleotidsequenz der Gen-Struktur calBΩGm
	FIG. 2f: Nukleotidsequenz der Gen-Struktur calBA
	FIG. 2g: Nukleotidsequenz der Gen-Struktur fcsΩKm
	FIG. 2h: Nukleotidsequenz der Gen-Struktur fcsΩGm
	FIG. 2i: Nukleotidsequenz der Gen-Struktur fcsΔ
10	FIG. 2j: Nukleotidsequenz der Gen-Struktur echΩKm
	FIG. 2k: Nukleotidsequenz der Gen-Struktur echΩGm
	FIG. 21: Nukleotidsequenz der Gen-Struktur ech∆
	FIG. 2m: Nukleotidsequenz der Gen-Struktur vdhΩKm
	FIG. 2n: Nukleotidsequenz der Gen-Struktur vdhΩGm
15	FIG. 20: Nukleotidsequenz der Gen-Struktur vdh∆
	FIG. 2p: Nukleotidsequenz der Gen-Struktur aatΩKm
	FIG. 2q: Nukleotidsequenz der Gen-Struktur aatΩGm
	FIG. 2r: Nukleotidsequenz der Gen-Struktur aatΔ

5

15

# Patentansprüche

- 1. Transformierter und/oder mutagenisierter ein- oder mehrzelliger Organismus, der dadurch gekennzeichnet ist, daß Enzyme des Eugenol- und/oder Ferulasäure-Katabolismus derart inaktiviert sind, daß eine Akkumulation der Intermediate Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure erfolgt.
- Organismus nach Anspruch 1, dadurch gekennzeichnet, daß der Eugenolund/oder Ferulasäure-Katabolismus durch Ω-Element-Insertion oder Einführen von Deletionen in entsprechende Gene verändert ist.
  - 3. Organismus nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß ein oder mehrere Gene, die für die Enzyme Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA Synthetasen, Enoyl-CoA Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen Enzyme kodieren, verändert und/oder inaktiviert sind.
- 4. Organismus nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß er einzellig, vorzugsweise ein Mikroorganismus oder eine pflanzliche oder eine tierische Zelle ist.
- 5. Organismus nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß er ein Bakterium, vorzugsweise eine *Pseudomonas*-Art ist.
- Gen-Strukturen, bei denen die für die Enzyme Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA Synthetasen, Enoyl-CoA Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen oder zweier oder mehrerer dieser Enzyme kodierenden Nukleotidsequenzen verändert und/oder inaktiviert sind.

5

15

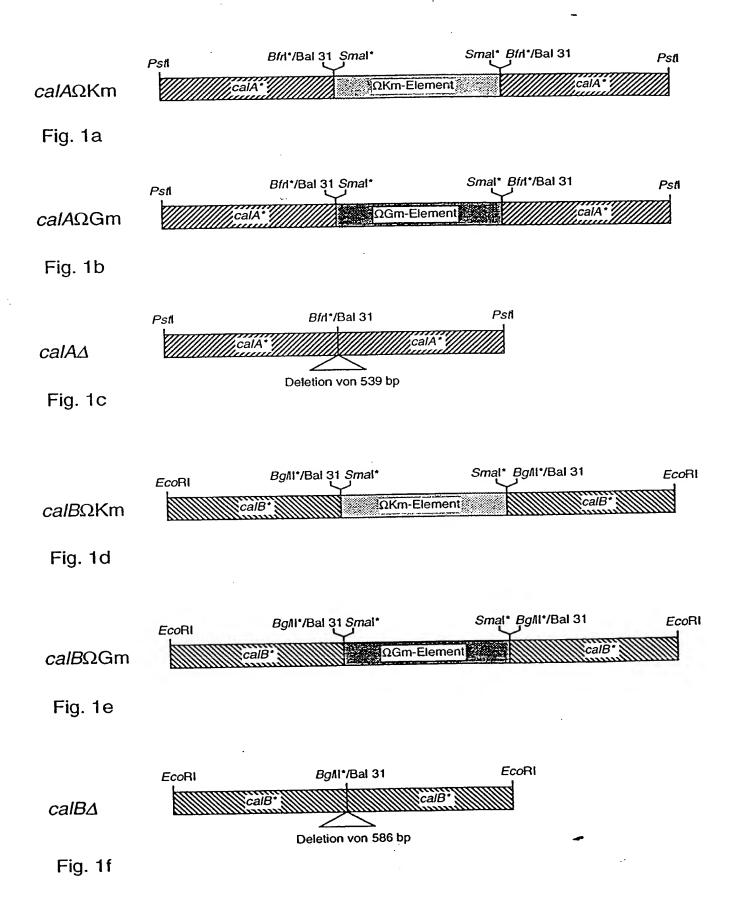
20

25

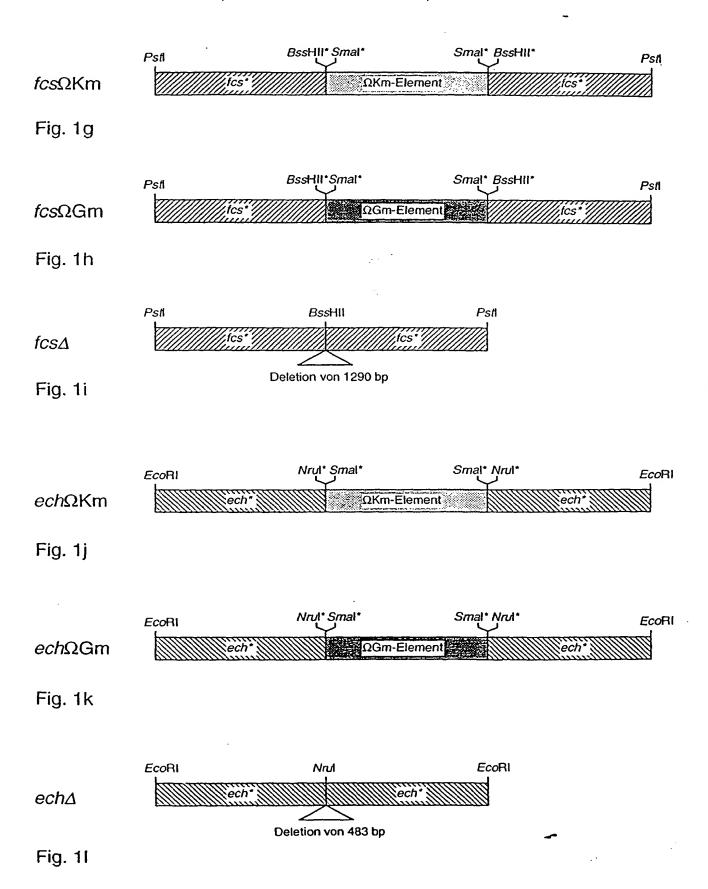
30

- 7. Gen-Strukturen mit den in Figur 1a bis 1r angegebenen Strukturen.
- 8. Gen-Strukturen mit den in Figur 2a bis 2r angegebenen Sequenzen.
- 9. Vektoren enthaltend wenigstens eine Gen-Struktur nach einem der Ansprüche 6 bis 8.
- Transformierter Organismus nach einem der Ansprüche 1 bis 5, dadurch
   gekennzeichnet, daß er wenigstens einen Vektor gemäß Anspruch 9 enthält.
  - 11. Organismus nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß er wenigstens eine Gen-Struktur nach einem der Ansprüche 6 bis 8 an Stelle des jeweiligen intakten Gens im Genom integriert enthält.
  - 12. Verfahren zur biotechnischen Herstellung von organischen Verbindungen, insbesondere von Alkoholen, Aldehyden und organischen Säuren, dadurch gekennzeichnet, daß ein Organismus nach einem der Ansprüche 1 bis 5 oder 10 bis 11 eingesetzt wird.
  - 13. Verfahren zur Herstellung der Organismen nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die Veränderung des Eugenol- und/oder Ferulasäure-Katabolismus mittels an sich bekannter mikrobiologischer Züchtungsmethoden erzielt wird.
  - 14. Verfahren zur Herstellung eines Organismus nach einem der Ansprüche 1 bis 5 oder 10 bis 11, dadurch gekennzeichnet, daß die Veränderung des Eugenolund/oder Ferulasäure-Katabolismus und/oder die Inaktivierung der entsprechenden Gene mittels gentechnischer Methoden erzielt wird.

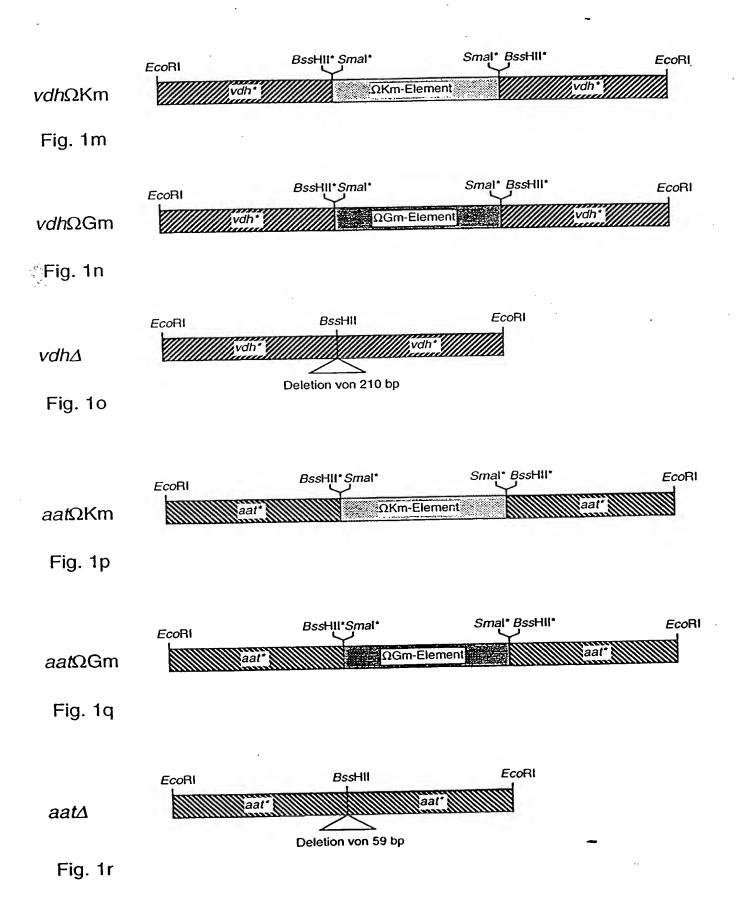
- 15. Verwendung der Organismen nach einem der Ansprüche 1 bis 5 oder 10 bis 11 zur Herstellung von Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure.
- 5 16. Verwendung von Gen-Strukturen nach einem der Ansprüche 6 bis 8 oder eines Vektors nach Anspruch 9 zur Herstellung transformierter und/oder mutagenisierter Organismen.



		Ç.		
			7	-
	÷			
				· ·
				2
	,			• 9
			_	



*			(*)	9,	-
**					
		\$			
		÷			
					, · ,
	7				
	*				
			•		
			•		



			-
er en			-,
			×
•			•
	,*** * <del>*</del>	•	

#### Sequenzen

CTGCAGCCAG GGCTGAAAAG GAGGGATTCA GTGAGGTCAT GAAGGGAGGG GACGGCGCCT	60
GGCTCCAATT GCTCGATGGC GCCGCGATTG AGTGTCTTGG GCGCGGTCTT GGAGAGTTCG	120
GCTAGGGAGA TAAATTTGCT GGCCATGGTG GCGGCCCCTG ATGGGTTGGA TGATTTTCTG	180
CATTCTGCAT CATGAAATTC ATGAAATCAT CACTTTTCGG GGGGTGGGTG CACGGGATTG	240
AAGGTTGCTA GGAGAGTGCA TTGCTCGTAA GCCCAGGAAG CACGCGGGTT TCAGGATGGT	300
GCATGGAAAT GGCATGAGCT TTGCTGGATA TGATTAGAGA CATTAACTAT TTTGGCGGAA	360
TGGAAGCACG ATTCCTCGCC CGGTAGAGCG GTAACCGCGA CATTCAGGAC CGTAAAAAGG	420
AAAGAGCATG CAA CTG ACC AAC AAG AAA ATC GTC GTC ACC GGA GTG TCC TCC  Met Gln Leu Thr Asn Lys Lys Ile Val Val Thr Gly Val Ser Ser  1 5 10 - 15	472
GGT ATC GGT GCC GAA ACT GCC CGC GTT CTG CGC TCT CAC GGC GCC ACA Gly Ile Gly Ala Glu Thr Ala Arg Val Leu Arg Ser His Gly Ala Thr 20 25 30	520
GTG ATT GGC GTA GAT CGC AAC ATG CCG AGC CTG ACT CTG GAT GCT TTC Val Ile Gly Val Asp Arg Asn Met Pro Ser Leu Thr Leu Asp Ala Phe 35 40 45	568
GTT CAG GCT GAC CTG AGC CAT CCT GAA GGC ATC GAT AAG GCC ATC GGG Val Gln Ala Asp Leu Ser His Pro Glu Gly Ile Asp Lys Ala Ile 50 55 60 62	616
ACAGCAAGCG AACCGGAATT GCCAGCTGGG GCGCCCTCTG GTAAGGTTGG GAAGCCCTGC	676
AAAGTAAACT GGATGGCTTT CTTGCCGCCA AGGATCTGAT GGCGCAGGGG ATCAAGATCT	736
GATCAAGAGA CAGGATGAGG ATCGTTTCGC ATG ATT GAA CAA GAT GGA TTG CAC Met Ile Glu Gln Asp Gly Leu His 1 5	790
GCA GGT TCT CCG GCC GCT TGG GTG GAG AGG CTA TTC GGC TAT GAC TGG Ala Gly Ser Pro Ala Ala Trp Val Glu Arg Leu Phe Gly Tyr Asp Trp 10 15 20	838
GCA CAA CAG ACA ATC GGC TGC TCT GAT GCC GCC GTG TTC CGG CTG TCA Ala Gln Gln Thr Ile Gly Cys Ser Asp Ala Ala Val Phe Arg Leu Ser 25 30 35 40	886
GCG CAG GGG CGC CCG GTT CTT TTT GTC AAG ACC GAC CTG TCC GGT GCC Ala Gln Gly Arg Pro Val Leu Phe Val Lys Thr Asp Leu Ser Gly Ala 45 50 55	934
CTG AAT GAA CTG CAG GAC GAG GCA GCG CGG CTA TCG TGG CTG GCC ACG Leu Asn Glu Leu Gln Asp Glu Ala Ala Arg Leu Ser Trp Leu Ala Thr 60 65 70	982

					-
				÷	
					-
				•	
					•
	1.			*** ****	
					3
	- 9				
	20				•
			-		
			-	-	

					GCA Ala										10	030
					TTG Leu										10	78
					GCC Ala 110										11	126
					CTT Leu										11	174
					GAG Glu										12	222
					CTG Leu										12	270
					CTC Leu										13	318
	Val				GAT Asp 190										13	366
					TTC Phe										14	414
				Ile	GCG Ala				Arg						1	462
			Trp		GAC Asp			Leu					Ile		15	510
		Ser					Phe					Asp		TTC Phe 264	1	558
ТGA	.GCGG	GAC	TCTG	GGGT	TC G	ТААА	GACC	G AC	CAAG	CGAC	GCC		GCC Ala 225		1	613
			Met					Ser					G1 y	AAT Asn	1	661

ès			-
			•
			*
	: .		
		*	
	-4		
		-	

ATT CCA GTG GAC GGA GGT TTG GCA TCG ACC TAC GTG TAA GTTCGTGGAC 1710 Ile Pro Val Asp Gly Gly Leu Ala Ser Thr Tyr Val 250 245 GCCCTTTGCA CGCGCACTAT ATCTCTATGC AGCAGCTGAA AGCAGCTTTG GTTTTGATCG 1770 GAGGTAGCGG GCGGAAAGGT GCAGAATGTC TAAATAATAA AGGATTCTTG TGAAGCTTTA 1830 GTTGTCCGTA AACGAAAATA AAAATAAAGA GGAATGATAT GAAAGCAAGT AGATCAGTCT 1890 GCACTTTCAA AATAGCTACC CTGGCAGGCG CCATTTATGC AGCGCTGCCA ATGTCAGCTG 1950 CAAACTCGAT GCAGCTGGAT GTAGGTAGCT CGGATTGGAC GGTGCGTTGG GGACAACACC 2010 CTCAAGTATA GCCTTGCCTC TCGCCTGAAT GAGCAAGACT CAAGTCTGAC AAATGCGCCG 2070 ACTGTCAATG GTTATATCCG GATATTCAAA GTCAGGGTGA TCGTAACTTT GACCGGGGGC 2130 2164 TTGGTATCCA ATCGTCTCGA TATTCTGGCT GCAG FIG. 2a:

					•
					,
				Ŷ	
					•
	<b>.</b>				
					•
	Ċ.				
			•		
			•		

CTGCAGCCAG GGCTGAAAAG GAGGGATTCA GTGAGGTCAT GAAGGGAGGG GACGGCGCCT	60
GGCTCCAATT GCTCGATGGC GCCGCGATTG AGTGTCTTGG GCGCGGTCTT GGAGAGTTCG	120
GCTAGGGAGA TAAATTTGCT GGCCATGGTG GCGGCCCCTG ATGGGTTGGA TGATTTTCTG	180
CATTCTGCAT CATGAAATTC ATGAAATCAT CACTTTTCGG GGGGTGGGTG CACGGGATTG	240
AAGGTTGCTA GGAGAGTGCA TTGCTCGTAA GCCCAGGAAG CACGCGGGTT TCAGGATGGT	300
GCATGGAAAT GGCATGAGCT TTGCTGGATA TGATTAGAGA CATTAACTAT TTTGGCGGAA	360
TGGAAGCACG ATTCCTCGCC CGGTAGAGCG GTAACCGCGA CATTCAGGAC CGTAAAAAGG	420
AAAGAGCATG CAA CTG ACC AAC AAG AAA ATC GTC GTC ACC GGA GTG TCC TCC  Met Gln Leu Thr Asn Lys Lys Ile Val Val Thr Gly Val Ser Ser  1 5 10 15	472
GGT ATC GGT GCC GAA ACT GCC CGC GTT CTG CGC TCT CAC GGC GCC ACA Gly Ile Gly Ala Glu Thr Ala Arg Val Leu Arg Ser His Gly Ala Thr 20 25 30	520
GTG ATT GGC GTA GAT CGC AAC ATG CCG AGC CTG ACT CTG GAT GCT TTC Val Ile Gly Val Asp Arg Asn Met Pro Ser Leu Thr Leu Asp Ala Phe 35 40 45	568
GTT CAG GCT GAC CTG AGC CAT CCT GAGGGGAGAG GCGGTTTGCG TATTGGGCGC Val Gln Ala Asp Leu Ser His Pro 50 55	622
ATGCATAAAA ACTGTTGTAA TTCATTAAGC ATTCTGCCGA CATGGAAGCC ATCACAAACG	682
GCATGATGAA CCTGAATCGC CAGCGGCATC AGCACCTTGT CGCCTTGCGT ATAATATTTG	742
CCCATGGACG CACACCGTGG AAACGGATGA AGGCACGAAC CCAGTTGACA TAAGCCTGTT	802
CGGTTCGTAA ACTGTAATGC AAGTAGCGTA TGCGCTCACG CAACTGGTCC AGAACCTTGA	862
CCGAACGCAG CGGTGGTAAC GGCGCAGTGG CGGTTTTCAT GGCTTGTTAT GACTGTTTTT	922
TTGTACAGTC TATGCCTCGG GCATCCAAGC AGCAAGCGCG TTACGCCGTG GGTCGATGTT	982
TGATGTTATG GAGCAGCAAC G ATG TTA CGC AGC AGC AAC GAT GTT ACG CAG Met Leu Arg Ser Ser Asn Asp Val Thr Gln 1 5 10	1033
CAG GGC AGT CGC CCT AAA ACA AAG TTA GGT GGC TCA AGT ATG GGC ATC Gln Gly Ser Arg Pro Lys Thr Lys Leu Gly Gly Ser Ser Met Gly Ile 15 20 25	1081
ATT CGC ACA TGT AGG CTC GGC CCT GAC CAA GTC AAA TCC ATG CGG GCT Ile Arg Thr Cys Arg Leu Gly Pro Asp Gln Val Lys Ser Met Arg Ala 30 35 40	1129
GCT CTT GAT CTT TTC GGT CGT GAG TTC GGA GAC GTA GCC ACC TAC TCC Ala Leu Asp Leu Phe Gly Arg Glu Phe Gly Asp Val Ala Thr Tyr Ser	1177

					- ·
					•
					·
					1.
		÷			
	÷				
		9			•
			4	, generalia	•
				e de la companya de l	

FIG. 2b:

												CTC				1225
												GTT Val				1273
												CGT Arg				1321
												AGG Arg				1369
												AAC Asn 135				1417
												GAT Asp				1465
												ATG Met				1513
		CCA Pro					TAA	CAA	rtcg1	TTC I	\AGC(	CGAGA	AT CO	GCTT	rccct	1567
I		la P					lu G					le A			TA AAT le Asn	1616
		GTG Val										ТАА	GTT	CGTG	GAC	1665
GCC	CTTT	GCA	CGCG	CACT	AT A	TCTC	TATG	C AG	CAGC'	rgaa	AGC	AGCT	rtg (	GTTT:	rga <b>t</b> cg	1725
GAG	GTAG	CGG	GCGG	AAAG	GT G	CAGA	ATGT	C TA	ATAA	AATA	AGG	ATTC:	rtg '	TGAA	GCTTTA	1785
GТТ	GTCC	GTA	AACG	AAAA	TA A	AAAT.	AAAG.	A GG	AATG	TATA	GAA	AGCA	AGT A	AGAT	CAGTCT	1845
GCA	СТТТ	CAA	ATAA	GCTA	cc c	TGGC.	AGGC	G CC.	ATTT	ATGC	AGC	GCTG	CCA	ATGT	CAGCTG	1905
CAA	ACTC	GAT	GCAG	CTGG	AT G	TAGG	TAGC	T CG	GATT	GGAC	GGT	GCGT	rgg (	GGAC	AACACC	1965
СТС	AAGT	ATA	GCCT	TGCC	TÇ T	CGCC	TGAA	T GA	GCAA	GACT	CAA	GTCT	GAC .	AAAT	GCGCCG	2025
ACT	GTCA	AТG	GTTA	TATC	CG G	АТАТ	тсаа	A GT	CAGG	GTGA	TCG	ТААС	TTT	GACC	GGGGGC	2085
ТТG	GTAT	CCA	ATCG	тстс	GA T	ATTC	TGGC	T GC	AG							2119
															_	

					ł.,	
						• . •
		*				
,	4.		÷			
						ī
	÷					
			• • • • • • • • • • • • • • • • • • •	٠.		

CTGCAGCCAG GGCTGAAAAG GAGGGATTCA GTGAGGTCAT GAAGGGAGGG GACGGCGCCT	60
GGCTCCAATT GCTCGATGGC GCCGCGATTG AGTGTCTTGG GCGCGGTCTT GGAGAGTTCG	120
GCTAGGGAGA TAAATTTGCT GGCCATGGTG GCGGCCCCTG ATGGGTTGGA TGATTTTCTG	180
CATTCTGCAT CATGAAATTC ATGAAATCAT CACTTTTCGG GGGGTGGGTG CACGGGATTG	240
AAGGTTGCTA GGAGAGTGCA TTGCTCGTAA GCCCAGGAAG CACGCGGGTT TCAGGATGGT	.300
GCATGGAAAT GGCATGAGCT TTGCTGGATA TGATTAGAGA CATTAACTAT TTTGGCGGAA	360
TGGAAGCACG ATTCCTCGCC CGGTAGAGCG GTAACCGCGA CATTCAGGAC CGTAAAAAGG	420
AAAGAGCATG CAA CTG ACC AAC AAG AAA ATC GTC GTC ACC GGA GTG TCC TCC  Met Gln Leu Thr Asn Lys Lys Ile Val Val Thr Gly Val Ser Ser  1 5 10 15	472
GGT ATC GGT GCC GAA ACT GCC CGC GTT CTG CGC TCT CAC GGC GCC ACA Gly Ile Gly Ala Glu Thr Ala Arg Val Leu Arg Ser His Gly Ala Thr 20 25 30	520
GTG ATT GGC GTA GAT CGC AAC ATG CCG AGC CTG ACT CTG GAT GCT TTC Val Ile Gly Val Asp Arg Asn Met Pro Ser Leu Thr Leu Asp Ala Phe 35	568
GTT CAG GCT GAC CTG AGC CAT CCT GAA GGC ATC GATC AAC GGC ATA AAT Val Gln Ala Asp Leu Ser His Pro Glu Gly Ile Asn Gly Ile Asn 50 55 58 240	617
ATT CCA GTG GAC GGA GGT TTG GCA TCG ACC TAC GTG TAA GTTCGTGGAC Ile Pro Val Asp Gly Gly Leu Ala Ser Thr Tyr Val 255 255	666
GCCCTTTGCA CGCGCACTAT ATCTCTATGC AGCAGCTGAA AGCAGCTTTG GTTTTGATCG	726
GAGGTAGCGG GCGGAAAGGT GCAGAATGTC TAAATAATAA AGGATTCTTG TGAAGCTTTA	786
GTTGTCCGTA AACGAAAATA AAAATAAAGA GGAATGATAT GAAAGCAAGT AGATCAGTCT	846
GCACTTTCAA AATAGCTACC CTGGCAGGCG CCATTTATGC AGCGCTGCCA ATGTCAGCTG	906
CAAACTCGAT GCAGCTGGAT GTAGGTAGCT CGGATTGGAC GGTGCGTTGG GGACAACACC	966
CTCAAGTATA GCCTTGCCTC TCGCCTGAAT GAGCAAGACT CAAGTCTGAC AAATGCGCCG	1026
ACTGTCAATG GTTATATCCG GATATTCAAA GTCAGGGTGA TCGTAACTTT GACCGGGGCC	1086
TTGGTATCCA ATCGTCTCGA TATTCTGGCT GCAG	1120
FIG. 2c:	

					٠.	•
						•
						-
•						
						,
						•
					(4)	
						,
						•
						•
				~		

GAATTCCGCG TA	CCCCCGG TT	CTATCAGC GGG	SCCGCTTT	CGAAAGTCAT G	GTGTTAGCC	60
GGTAGGGTCT TT	ITCTTGGC CA	TGCTTGTT GCC	CTGAACCT	TCGTTGACAT A	GGGCAGAGG	120
TGCGTTTGCC GC	TTCGCTTC GC	GATGAACC GCA	ATCGAGAT	GCTGAGGTCA G	GATTTTTCC	180
TTAACTCGCG TA	AGCATTCT GT	CATTTTTT TGO	STGGCTTT	GAACAGCCTG A	TGAAAGGTG	240
GTCTCGCCCT TT	GAGGCCGA TT	CTTGGGCG CT	rggcggcg	TCGAAGCGAT G	CTCCACTAC	.300
CGATTAAGAT AA	TTAAAATA AG	GAAACCGC ATC	GGTTTCTT	ATGTGAATTT G	STCTGGCATA	360
CTCCAGCTCA AG	GGCAATTT TT	GGGCTATT GGG	CTGAGCAG	TTGCCTCTAT A	TGGTTATTC	420
AGAATAACAA TT	GACTCCTC AG	GAGGTCAG CG		ATT CTT GGT Ile Leu Gly 5		473
GGT GCC CCG G						521
Gly Ala Pro V 10	ar Gry Ara	15	Gry Ser	20	Arg Met	
AAG AAG GCG C Lys Lys Ala H				<del>-</del>		569
25	13 Ded Glu	30	Ald Asii	35	Arg Lea	
AGT AGG CTG G Ser Arg Leu A	· - ·			<del>-</del>		617
40	45	110 /110 //00	50	014 1.5.7 1.19	55	
ATT GCC GAC G Ile Ala Asp A						665
•	60	-	65		70	
ACA CTG CTT T Thr Leu Leu C						713
	75	80		85		
CGC GAG CAC G Arg Glu His V						761
90		95		100		
TTT CCA GGG G		Arg Val Glu		Pro Leu Gly		809
105		110		115		
GGG GTC ATT A	Ser Pro Trp		Ile Val		Gly Pro	857
120	125		130		135	
CTG GCC GGC F Leu Ala Gly I	le Phe Ala		Arg Ala		Pro Ser	905
	140	non os- s	145	<b>0.5.0</b> 00= 0.00	150	
GAG CTT ACC C Glu Leu Thr I			Leu Ala			953

						_
			3.43			
						•
					•	
					•	
						•
					1	
	*					
<b>9</b> 3						
						•
						•
÷.		 •		_		
				•		

·	
TAC TTC GAT GAA ACT GAG CTG ACT ACA GTG CTG GGC GAC GCT GA Tyr Phe Asp Glu Thr Glu Leu Thr Thr Val Leu Gly Asp Ala Gl 170 175 180	
GGT GCG CTG TTC AGT GCT CAG CCT TTC GAT CAT CTG ATC TTC ACGLY Ala Leu Phe Ser Ala Gln Pro Phe Asp His Leu Ile Phe Th 185	
GGC ACT GCC GTG GCC AAG CAC ATC ATG CGT GCC GCG GCG GAT AAG Gly Thr Ala Val Ala Lys His Ile Met Arg Ala Ala Ala Asp As 200 205 210	
GTG CCC GTT ACC CTG GAA TTG GGT GGC AAA TCG CCG GTG ATC GT Val Pro Val Thr Leu Glu Leu Gly Gly Lys Ser Pro Val Ile Va 220 225 23	l Ser
CGC AGT GCA GAT ATG GCG GAC GTT GCA CAA CGG GTG TTG ACG GT Arg Ser Ala Asp Met Ala Asp Val Ala Gln Arg Val Leu Thr Va 235 240 245	
ACC TTC AAT GCC GGG CAA ATC TGT CTG GCA CCG GAC TAT GTG CT Thr Phe Asn Ala Gly Gln Ile Cys Leu Ala Pro Asp Tyr Val Le 250 255 260	
CCG GAA GGGACAGCAA GCGAACCGGA ATTGCCAGCT GGGGCGCCCT CTGGTA Pro Glu 265	AGGT 1297
TGGGAAGCCC TGCAAAGTAA ACTGGATGGC TTTCTTGCCG CCAAGGATCT GAT	GGCGCAG 1357
GGGATCAAGA TCTGATCAAG AGACAGGATG AGGATCGTTT CGC ATG ATT GA Met Ile Gl	
GAT GGA TTG CAC GCA GGT TCT CCG GCC GCT TGG GTG GAG AGG CT Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val Glu Arg Le 5 10 15	
GGC TAT GAC TGG GCA CAA CAG ACA ATC GGC TGC TCT GAT GCC GC Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser Asp Ala Al 25 30 3	
TTC CGG CTG TCA GCG CAG GGG CGC CCG GTT CTT TTT GTC AAG AC Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe Val Lys Th	
CTG TCC GGT GCC CTG AAT GAA CTG CAG GAC GAG GCA GCG CGG CT Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala Ala Arg Le 55 60 65	

				*	
					-
				, ,	,
					4.
		·			
	·				•
					٠.
			•		
10					

			GGA Gly													1700
			TCA Ser													1748
			ATG Met 120													1796
			CAA Gln												_	1844
			CTT Leu													1892
			GCC Ala													1940
			CTC Leu	-												1988
			AAT Asn 200													2036
			GAC Asp													2084
			CTT Leu													2132
	Ile		GCT Ala													2180
			TTC Phe 264		GCG	GGAC	TCT	GGGG	TTCG	'A AA	TGAC	CGAC	C AA	GCGA	CGCC	2235
CGC	His		Lys					Gln					Glu		AAC Asn	2283
		Met					Gly					Gly			TCT Ser	2331

			, i		
		24			
\$.					
				Ŕ	
					(Agr
		3.3			
		·	•		

GTC CTC CTT TCA ACG GAG TGT TAG AACCGTTGGT AGTGGTTTTG GACGGGCCCA Val Leu Ser Thr Glu Cys 475 480 481	2385
GGAGCATGCG CTTCTGGGCC CGTTTCTTGA GTATTCATTG GATAGTCACG CGTGGTAGCT	2445
TCGAGCCTGC ACAGCTGATG AGCACCCTGG AAGGCGCGCT GTACGCGGAC GACTGGGTTC	2505
ATCTTCGCCA TTCATGACGG AACTCCGTTC CCCAGTACCG CGATGACTAT TTTGCCTCTT	2565
CCGATGTCCG ATTCCACGCC GCCTGACGCT AAGCGGGGGC GGGGGCGCCC GCATCCCAGC	2625
CCAGACAGCA ACAAATGAGT AGGCTCTTGG ATGCCGCGGC GGCTGAGATT GGTAACGGCA	2685
ATTTCGTCAA TGTGACGATG GATTCGATTG CCCGTGCTGC CGGCGTCTCA AAAAAAAACGC	2745
TGTACGTCTT GGTGGCGAGC AAGGAAGAAC TCATTTCCCG GTTAGTGGCT CGAGACATGT	2805
CCAACCTTGA GGAATTC	2822
FIG 2d.	

FIG. 2d:

					+
			,		
÷					
					( <b></b> )
		Ţ			
	· · ·			•	

			•	
GAATTCCGCG TATCG	CCCGG TTCTATCAG	C GGGCCGCTTT	CGAAAGTCAT GGTGTTAGCC	60
GGTAGGGTCT TTTTC	TTGGC CATGCTTGT	T GCCTGAACCT	TCGTTGACAT AGGGCAGAGG	120
TGCGTTTGCC GCTTC	GCTTC GCGATGAAC	C GCATCGAGAT	GCTGAGGTCA GGATTTTTCC	180
TTAACTCGCG TAAGC	ATTCT GTCATTTT	T TGGTGGCTTT	GAACAGCCTG ATGAAAGGTG	240
GTCTCGCCCT TTGAG	GCCGA TTCTTGGGC	C CTTGGCGGCG	TCGAAGCGAT GCTCCACTAC	300
CGATTAAGAT AATTA	AAATA AGGAAACCG	C ATGGTTTCTT	ATGTGAATTT GTCTGGCATA	360
CTCCAGCTCA AGGGC	AATTT TTGGGCTAT	T GGCTGAGCAG	TTGCCTCTAT ATGGTTATTC	420
AGAATAACAA TTGAC	TCCTC AGGAGGTCA		ATT CTT GGT TTG AAT Ile Leu Gly Leu Asn 5	473
		Leu Gly Ser	GCT CTT GAT CGC ATG Ala Leu Asp Arg Met 20	521
			TTG GAG CTG CGT CTG Leu Glu Leu Arg Leu 35	569
			GAA AAT CGT GAA GCA Glu Asn Arg Glu Ala 55	617
			CGC AGC CGT GAG CAA Arg Ser Arg Glu Gln 70	665
			AGC CTG AAG GAT AGC Ser Leu Lys Asp Ser 85	713
		Glu Pro Glu	CAT CAC AAG GCG ATG His His Lys Ala Met 100	761
			CCG CTG GGT GTC GTT Pro Leu Gly Val Val 115	809
			CTG GCC TTT GGG CCG Leu Ala Phe Gly Pro 135	857
			ATG CTC AAG CCG TCC Met Leu Lys Pro Ser 150	905
			GAG CTA ATT GCT CGT Glu Leu Ile Ala Arg 165	953

			,
			<u>.</u>
			2
			· .
			Ţ.
			₹,
	**		
- 1			
		~	
		**	

Tyr Phe Asp Glu Thr Glu Leu Thr Thr Val Leu Gly Asp Ala Glu Val 170 175 180	1001
GGT GCG CTG TTC AGT GCT CAG CCT TTC GAT CAT CTG ATC TTC ACC GGC Gly Ala Leu Phe Ser Ala Gln Pro Phe Asp His Leu Ile Phe Thr Gly 185	1049
GGC ACT GCC GTG GCC AAG CAC ATC ATG CGT GCC GCG GCG GAT AAC CTA Gly Thr Ala Val Ala Lys His Ile Met Arg Ala Ala Ala Asp Asn Leu 200 205 210 215	1097
GTG CCC GTT ACC CTG GAA TTG GGT GGC AAA TCG CCG GTG ATC GTT TCC Val Pro Val Thr Leu Glu Leu Gly Gly Lys Ser Pro Val Ile Val Ser 220 230	1145
CGC AGT GCA GAT ATG GCG GAC GTT GCA CAA CGG GTG TTG ACG GTG AAA Arg Ser Ala Asp Met Ala Asp Val Ala Gln Arg Val Leu Thr Val Lys 235 240 245	1193
ACC TTC AAT GCC GGG CAA ATC TGT CTG GCA CCG GAC TAT GTG CTG GGG Thr Phe Asn Ala Gly Gln Ile Cys Leu Ala Pro Asp Tyr Val Leu 250 255 260 262	1241
GAGAGGCGGT TTGCGTATTG GGCGCATGCA TAAAAACTGT TGTAATTCAT TAAGCATTCT	1301
GCCGACATGG AAGCCATCAC AAACGGCATG ATGAACCTGA ATCGCCAGCG GCATCAGCAC	1361
CTTGTCGCCT TGCGTATAAT ATTTGCCCAT GGACGCACAC CGTGGAAACG GATGAAGGCA	1421
CGAACCCAGT TGACATAAGC CTGTTCGGTT CGTAAACTGT AATGCAAGTA GCGTATGCGC	1481
TCACGCAACT GGTCCAGAAC CTTGACCGAA CGCAGCGGTG GTAACGGCGC AGTGGCGGTT	1541
TTCATGGCTT GTTATGACTG TTTTTTTGTA CAGTCTATGC CTCGGGCATC CAAGCAGCAA	1601
GCGCGTTACG CCGTGGGTCG ATGTTTGATG TTATGGAGCA GCAACG ATG TTA CGC  Met Leu Arg  1	1601 1656
GCGCGTTACG CCGTGGGTCG ATGTTTGATG TTATGGAGCA GCAACG ATG TTA CGC Met Leu Arg	
GCGCGTTACG CCGTGGGTCG ATGTTTGATG TTATGGAGCA GCAACG ATG TTA CGC Met Leu Arg  1  AGC AGC AAC GAT GTT ACG CAG CAG GGC AGT CGC CCT AAA ACA AAG TTA Ser Ser Asn Asp Val Thr Gln Gln Gly Ser Arg Pro Lys Thr Lys Leu	1656
GCGCGTTACG CCGTGGGTCG ATGTTTGATG TTATGGAGCA GCAACG ATG TTA CGC Met Leu Arg 1  AGC AGC AAC GAT GTT ACG CAG CAG GGC AGT CGC CCT AAA ACA AAG TTA Ser Ser Asn Asp Val Thr Gln Gln Gly Ser Arg Pro Lys Thr Lys Leu 5  10  15  GGT GGC TCA AGT ATG GGC ATC ATT CGC ACA TGT AGG CTC GGC CCT GAC Gly Gly Ser Ser Met Gly Ile Ile Arg Thr Cys Arg Leu Gly Pro Asp	1656 1704
GCGCGTTACG CCGTGGGTCG ATGTTTGATG TTATGGAGCA GCAACG ATG TTA CGC Met Leu Arg  1  AGC AGC AAC GAT GTT ACG CAG CAG GGC AGT CGC CCT AAA ACA AAG TTA Ser Ser Asn Asp Val Thr Gln Gln Gly Ser Arg Pro Lys Thr Lys Leu 5  GGT GGC TCA AGT ATG GGC ATC ATT CGC ACA TGT AGG CTC GGC CCT GAC Gly Gly Ser Ser Met Gly Ile Ile Arg Thr Cys Arg Leu Gly Pro Asp 20  CAA GTC AAA TCC ATG CGG GCT GCT CTT GAT CTT TTC GGT CGT GAG TTC Gln Val Lys Ser Met Arg Ala Ala Leu Asp Leu Phe Gly Arg Glu Phe	1656 1704 1752

	•		
			•
			J.
		*	
•			
		~	

CAA GAA GCG GTT GTT GGC GCT CTC GCG GCT TAC GTT CTG CCC AGG TTT Gln Glu Ala Val Val Gly Ala Leu Ala Ala Tyr Val Leu Pro Arg Phe 85 90 95	1944
GAG CAG CCG CGT AGT GAG ATC TAT ATC TAT GAT CTC GCA GTC TCC GGC Glu Gln Pro Arg Ser Glu Ile Tyr Ile Tyr Asp Leu Ala Val Ser Gly 100 105 110	1992
GAG CAC CGG AGG CAG GGC ATT GCC ACC GCG CTC ATC AAT CTC CTC AAG Glu His Arg Arg Gln Gly Ile Ala Thr Ala Leu Ile Asn Leu Lys 120 125 130	2040
CAT GAG GCC AAC GCG CTT GGT GCT TAT GTG ATC TAC GTG CAA GCA GAT His Glu Ala Asn Ala Leu Gly Ala Tyr Val Ile Tyr Val Gln Ala Asp 135 140 145	2088
TAC GGT GAC GAT CCC GCA GTG GCT CTC TAT ACA AAG TTG GGC ATA CGG Tyr Gly Asp Asp Pro Ala Val Ala Leu Tyr Thr Lys Leu Gly Ile Arg 150 155 160	2136
GAA GAA GTG ATG CAC TTT GAT ATC GAC CCA AGT ACC GCC ACC TAA CAA Glu Glu Val Met His Phe Asp Ile Asp Pro Ser Thr Ala Thr 165 170 175 177	2184
TTCGTTCAAG CCGAGATCGG CTTCCCTG CAA AGT CCT GTG GGT GAG TCG AAC Gln Ser Pro Val Gly Glu Ser Asn 451 455	2236
TTG GCG ATG CGC GCA CCC TAC GGA GAA GCG ATC CAC GGA CTG CTC TCT Leu Ala Met Arg Ala Pro Tyr Gly Glu Ala Ile His Gly Leu Leu Ser 460 465 470	2284
GTC CTC CTT TCA ACG GAG TGT TAG AACCGTTGGT AGTGGTTTTG GACGGGCCCA Val Leu Leu Ser Thr Glu Cys 475 480 481	2338
GGAGCATGCG CTTCTGGGCC CGTTTCTTGA GTATTCATTG GATAGTCACG CGTGGTAGCT	2398
TCGAGCCTGC ACAGCTGATG AGCACCCTGG AAGGCGCGCT GTACGCGGAC GACTGGGTTC	2458
ATCTTCGCCA TTCATGACGG AACTCCGTTC CCCAGTACCG CGATGACTAT TTTGCCTCTT	2518
CCGATGTCCG ATTCCACGCC GCCTGACGCT AAGCGGGGGC GGGGGCGCCC GCATCCCAGC	2578
CCAGACAGCA ACAAATGAGT AGGCTCTTGG ATGCCGCGGC GGCTGAGATT GGTAACGGCA	2638
ATTTCGTCAA TGTGACGATG GATTCGATTG CCCGTGCTGC CGGCGTCTCA AAAAAAACGC	2698
TGTACGTCTT GGTGGCGAGC AAGGAAGAAC TCATTTCCCG GTTAGTGGCT CGAGACATGT	2758
CCAACCTTGA GGAATTC	2775

FIG. 2e:

		ý.			· · · · · · · · · · · · · · · · · · ·
	*				
					-
			1		÷
					٠
					•
					÷.4.
17					
					·
;		S.			
					•
				•	
					n 1 1 1

GAATTCCGCG TATCGCCC	GG TTCTATCAGO	C GGGCCGCTTT	CGAAAGTCAT GGTGTTAGCC	60
GGTAGGGTCT TTTTCTTG	GC CATGCTTGT	r gcctgaacct	TCGTTGACAT AGGGCAGAGG	120
TGCGTTTGCC GCTTCGCT	rc gcgatgaac	C GCATCGAGAT	GCTGAGGTCA GGATTTTTCC	180
TTAACTCGCG TAAGCATT	CT GTCATTTTT	TGGTGGCTTT	GAACAGCCTG ATGAAAGGTG	240
GTCTCGCCCT TTGAGGCC	GA TTCTTGGGC	G CTTGGCGGCG	TCGAAGCGAT GCTCCACTAC	300
CGATTAAGAT AATTAAAA	TA AGGAAACCG	C ATGGTTTCTT	ATGTGAATTT GTCTGGCATA	360
CTCCAGCTCA AGGGCAAT	TT TTGGGCTAT	r ggctgagcag	TTGCCTCTAT ATGGTTATTC	420
AGAATAACAA TTGACTCC	TC AGGAGGTCA		ATT CTT GGT TTG AAT Ile Leu Gly Leu Asn 5	473
			GCT CTT GAT CGC ATG Ala Leu Asp Arg Met 20	521
			TTG GAG CTG CGT CTG Leu Glu Leu Arg Leu 35	569
			GAA AAT CGT GAA GCA Glu Asn Arg Glu Ala 55	617
	Ser Ala Asp		CGC AGC CGT GAG CAA Arg Ser Arg Glu Gln 70	665
			AGC CTG AAG GAT AGC Ser Leu Lys Asp Ser 85	713
		Glu Pro Glu	CAT CAC AAG GCG ATG His His Lys Ala Met 100	761
			CCG CTG GGT GTC GTT Pro Leu Gly Val Val 115	809
			CTG GCC TTT GGG CCG Leu Ala Phe Gly Pro 135	857
	e Ala Ala Gly		ATG CTC AAG CCG TCC Met Leu Lys Pro Ser 150	905
			GAG CTA ATT GCT CGT Glu Leu Ile Ala Arg 165 -	953

1. 2	•		
	÷		
			( <del>-</del> )
			÷
			.*-
4			
			·
		~	

TAC TTC GAT GAA ACT GAG CTG ACT ACA GTG CTG GGC GAC GCT GAA GTC Tyr Phe Asp Glu Thr Glu Leu Thr Thr Val Leu Gly Asp Ala Glu Val 170 175 180	1001
GGT GCG CTG TTC AGT GCT CAG CCT TTC GAT CAT CTG ATC TTC ACC GGC Gly Ala Leu Phe Ser Ala Gln Pro Phe Asp His Leu Ile Phe Thr Gly 185	1049
GGC ACT GCC GTG GCC AAG CAC ATC ATG CGT GCC GCG GCG GAT AAC CTA Gly Thr Ala Val Ala Lys His Ile Met Arg Ala Ala Ala Asp Asn Leu 200 205 210 215	1097
GTG CCC GTT ACC CTG GAA TTG GGT GGC AAA TCG CCG GTG ATC GTT TCC Val Pro Val Thr Leu Glu Leu Gly Gly Lys Ser Pro Val Ile Val Ser 220 225 230	1145
CGC AGT GCA GAT ATG GCG GAC GTT GCA CAA CGG GTG TTG ACG GTG AAA Arg Ser Ala Asp Met Ala Asp Val Ala Gln Arg Val Leu Thr Val Lys 235 240 245	1193
ACC TTC AAT GCC GGG CAA ATC TGT CTG GCA CC GTG GGT GAG TCG AAC Thr Phe Asn Ala Gly Gln Ile Cys Leu Ala Val Gly Glu Ser Asn 250 255 257 454 455	1240
TTG GCG ATG CGC GCA CCC TAC GGA GAA GCG ATC CAC GGA CTG CTC TCT Leu Ala Met Arg Ala Pro Tyr Gly Glu Ala Ile His Gly Leu Leu Ser 460 465 470	1288
GTC CTC CTT TCA ACG GAG TGT TAG AACCGTTGGT AGTGGTTTTG GACGGGCCCA Val Leu Leu Ser Thr Glu Cys 475 480 481	1342
GGAGCATGCG CTTCTGGGCC CGTTTCTTGA GTATTCATTG GATAGTCACG CGTGGTAGCT	1402
TCGAGCCTGC ACAGCTGATG AGCACCCTGG AAGGCGCGCT GTACGCGGAC GACTGGGTTC	1462
ATCTTCGCCA TTCATGACGG AACTCCGTTC CCCAGTACCG CGATGACTAT TTTGCCTCTT	1522
CCGATGTCCG ATTCCACGCC GCCTGACGCT AAGCGGGGGC GGGGGCGCCC GCATCCCAGC	1582
CCAGACAGCA ACAAATGAGT AGGCTCTTGG ATGCCGCGGC GGCTGAGATT GGTAACGGCA	1642
ATTTCGTCAA TGTGACGATG GATTCGATTG CCCGTGCTGC CGGCGTCTCA AAAAAAACGC	1702
TGTACGTCTT GGTGGCGAGC AAGGAAGAAC TCATTTCCCG GTTAGTGGCT CGAGACATGT	1762
CCAACCTTGA GGAATTC	1779

FIG. 2f:

			ě	
			ű.	
				-
	ŝ,		4.0	
				٠
		\$-		
		•		
		<b>(</b> )		
				2.0
•				
				•
		~		
			*.	

			·	
CTGCAGCCGA GCAT	CGATTG AGCACT	TTAC CCAGCTGC	GC TGGCTGACCA TTCAGAATGG	60
CCCGCGGCAC TATC	CAATCT AAATCO	SATCT TCGGGCGC	CG CGGGCATCAT GCCCGCGGCG	120
CTCGCCTCAT TTCA	ATCTCT AACTTC	SATAA AAACAGAG	CT GTTCTCCGGT CTTGGTGGAT	180
CAAGGCCAGT CGCG	GAGAGT CTCGA	AGAGG AGAGTACA	GT GAACGCCGAG TCCACATTGC	240
AACCGCAGGC ATCA	TCATGC TCTGCT	CAGC CACGCTAC	CG CAGTGTGTCG ATTGGTCATC	-300
CTCCGGTTGA GGTT	ACGCAA GACGCT	rggag gtattgtc	CG G ATG CGT TCT CTC GAG  Met Arg Ser Leu Glu  1 5	356
			AG CGT CTC GAG CAT TGG lu Arg Leu Glu His Trp 20	404
	Pro Glu Gln		CT GCC AGG GCG GCA AAT la Ala Arg Ala Ala Asn 35	452
			TG TTC CAC AAC GTC CGC et Phe His Asn Val Arg 50	500
			TA TCG GCA GAG CGT CCG eu Ser Ala Glu Arg Pro 65	548
			AT CTT CAG CTG GCA TTT is Leu Gln Leu Ala Phe 80 85	596
			CG GTG TCT CCT GCT TAT ro Val Ser Pro Ala Tyr 100	644
	Gln Asp Leu		GT CAC ATC GTA GGT CTT arg His Ile Val Gly Leu 115	692
0.0 0.1. 0.0 00.		• • • • • • • • • • • • • • • • • • • •	CCA GCA CCT TTC CAG GGG Ala Ala Pro Phe Gln 130 132	740
ACAGCAAGCG AAC	CGGAATT GCCAG	CTGGG GCGCCCTC	TTG GTAAGGTTGG GAAGCCCTGC	800
AAAGTAAACT GGA	rggcttt cttgc	CGCCA AGGATCTO	SAT GGCGCAGGGG ATCAAGATCT	860
GATCAAGAGA CAG	GATGAGG ATCGT		GAA CAA GAT GGA TTG CAC Glu Gln Asp Gly Leu His S	914
		Val Glu Arg I	CTA TTC GGC TAT GAC TGG Leu Phe Gly Tyr Asp Trp 20	962

		•	
			•
	4		
			40
÷		_	

							GTG Val			1010
							GAC Asp			1058
							TCG Ser			1106
							GTC Val			1154
							CAG Gln 100			1202
							ATG Met			1250
							CCA Pro			1298
							ATG Met			1346
							GGG Gly			1394
							GAC Asp 180			1442
							ATC Ile			1490
			Phe				CTG Leu		 	1538
		Ile					ATT Ile			1586
	Trp				Leu		TAC Tyr			1634

			-
			-
			•
			÷
			•
	· · · · · ·		
	74.		
3,		(4) (*)	
		(2) 1	
			•
•			
			•
		-	
		** <b>-</b>	

CCC GAT TCG CAG CGC ATC GCC TTC TAT CGC CTT CTT GAC GAG TTC TTC Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu Asp Glu Phe Phe 250 264	1682
TGAGCGGGAC TCTGGGGTTC GAAATGACCG ACCAAGCGAC GCCCCT GTT TTG CAA Val Leu Gln 563 565	1737
TGG CGG TCG GCG AAA GTT GAT GCG CTG TAT CGT GGT GAA GAT CAA TCC Trp Arg Ser Ala Lys Val Asp Ala Leu Tyr Arg Gly Glu Asp Gln Ser 570 575 580	1785
ATG CTG CGT GAC GAG GCC ACA CTG TGA GTTGGTCAGG GGGGGCTTAC  Met Leu Arg Asp Glu Ala Thr Leu  585 589	1832
TCGGCGTTTT CCGACACTGC GTTGGTTGCG GCAGTGCGCA CCCCCTGGAT TGATTGCGGG	1892
GGTGCCCTGT CGCTGGTGTC GCCTATCGAC TTAGGGGTAA AGGTCGCTCG CGAAGTTCTG	1952
ATGCGTGCGT CGCTTGAACC ACAAATGGTC GATAGCGTAC TCGCAGGCTC TATGGCTCAA	2012
GCAAGCTTTG ATGCTTACCT GCTCCCGCGG CACATTGGCT TGTACAGCGG TGTTCCCAAG	2072
TCGGTTCCGG CCTTGGGGGT GCAGCGCATT TGCGGCACAG GCTTCGAACT GCTTCGGCAG	2132
GCCGGCGAGC AGATTTCCCA AGGCGCTGAT CACGTGCTGT GTGTCGCGGG CTGCAG	2188
FIG. 2g:	

÷					
					-
					3
					•
					•
					4, 4
	*				
					•
		į.			
				~	
				.,	

·	
CTGCAGCCGA GCATCGATTG AGCACTTTAC CCAGCTGCGC TGGCTGACCA TTCAGAATGG	60
CCCGCGGCAC TATCCAATCT AAATCGATCT TCGGGCGCCG CGGGCATCAT GCCCGCGGCG	120
CTCGCCTCAT TTCAATCTCT AACTTGATAA AAACAGAGCT GTTCTCCGGT CTTGGTGGAT	180
CAAGGCCAGT CGCGGAGAGT CTCGAAGAGG AGAGTACAGT GAACGCCGAG TCCACATTGC	240
AACCGCAGGC ATCATCATGC TCTGCTCAGC CACGCTACCG CAGTGTGTCG ATTGGTCATC	300
CTCCGGTTGA GGTTACGCAA GACGCTGGAG GTATTGTCCG G ATG CGT TCT CTC GAG Met Arg Ser Leu Glu 1 5	356
GCG CTT CTT CCC TTC CCG GGT CGA ATT CTT GAG CGT CTC GAG CAT TGG Ala Leu Leu Pro Phe Pro Gly Arg Ile Leu Glu Arg Leu Glu His Trp  10 15 20	404
GCT AAG ACC CGT CCA GAA CAA ACC TGC GTT GCT GCC AGG GCG GCA AAT Ala Lys Thr Arg Pro Glu Gln Thr Cys Val Ala Ala Arg Ala Ala Asn 25 30 35	452
GGG GAA TGG CGT CGT ATC AGC TAC GCG GAA ATG TTC CAC AAC GTC CGC Gly Glu Trp Arg Arg Ile Ser Tyr Ala Glu Met Phe His Asn Val Arg 40 45 50	500
GCC ATC GCA CAG AGC TTG CTT CCT TAC GGA CTA TCG GCA GAG CGT CCG Ala Ile Ala Gln Ser Leu Leu Pro Tyr Gly Leu Ser Ala Glu Arg Pro 55 60 65	548
CTG CTT ATC GTC TCT GGA AAT GAC CTG GAA CAT CTT CAG CTG GCA TTT Leu Leu Ile Val Ser Gly Asn Asp Leu Glu His Leu Gln Leu Ala Phe 70 75 80 85	596
GGG GCT ATG TAT GCG GGC ATT CCC TAT TGC CCG GTG TCT CCT GCT TAT Gly Ala Met Tyr Ala Gly Ile Pro Tyr Cys Pro Val Ser Pro Ala Tyr 90 95 100	644
TCA CTG CTG TCG CAA GAT TTG GCG AAG CTG CGT CAC ATC GTA GGT CTT Ser Leu Leu Ser Gln Asp Leu Ala Lys Leu Arg His Ile Val Gly Leu 105 110 115	692
CTG CAA CCG GGA CTG GTC TTT GCT GCC GAT GCA GCA CCT TTC CAG GGG Leu Gln Pro Gly Leu Val Phe Ala Ala Asp Ala Ala Pro Phe Gln 120 125 130 132	740
GAGAGGCGGT TTGCGTATTG GGCGCATGCA TAAAAACTGT TGTAATTCAT TAAGCATTCT	800
GCCGACATGG AAGCCATCAC AAACGGCATG ATGAACCTGA ATCGCCAGCG GCATCAGCAC	860
CTTGTCGCCT TGCGTATAAT ATTTGCCCAT GGACGCACAC CGTGGAAACG GATGAAGGCA	920
CGAACCCAGT TGACATAAGC CTGTTCGGTT CGTAAACTGT AATGCAAGTA GCGTATGCGC	980
TCACGCAACT GGTCCAGAAC CTTGACCGAA CGCAGCGGTG GTAACGGCGC AGTGGCGGTT	1040
TTCATGGCTT GTTATGACTG TTTTTTTGTA CAGTCTATGC CTCGGGCATC CAAGCAGCAA	1100

	e.		·.
			-
•			
	*		
•			•
		•	
		<i>)</i> .	

GCGCGTTACG CCGTC	GGTCG ATGTTTGA	ATG TTATGGAGCA	GCAACG ATG TTA CGC Met Leu Arg 1	1155
			CCT AAA ACA AAG TTA Pro Lys Thr Lys Leu 15	1203
			AGG CTC GGC CCT GAC Arg Leu Gly Pro Asp 35	1251
			TTC GGT CGT GAG TTC Phe Gly Arg Glu Phe 50	1299
			GAC TCC GAT TAC CTC Asp Ser Asp Tyr Leu 65	1347
	Arg Ser Lys Th		CTT GCT GCC TTC GAC Leu Ala Ala Phe Asp 80	1395
			GTT CTG CCC AGG TTT Val Leu Pro Arg Phe 95	1443
			CTC GCA GTC TCC GGC Leu Ala Val Ser Gly 115	1491
			ATC AAT CTC CTC AAG Ile Asn Leu Leu Lys 130	1539
			TAC GTG CAA GCA GAT Tyr Val Gln Ala Asp 145	1587
Tyr Gly Asp Asp		la Leu Tyr Thr	AAG TTG GGC ATA CGG Lys Leu Gly Ile Arg 160	1635
GAA GAA GTG ATG Glu Glu Val Met 165			ACC GCC ACC TAA CAA Thr Ala Thr 175 177	1683
TTCGTTCAAG CCGA	GATCGG CTTCCCC		TGG CGG TCG GCG AAA Trp Arg Ser Ala Lys 570	1735
			ATG CTG CGT GAC GAG Met Leu Arg Asp Glu 585	1783

é ş				; ····································
				• 7
				•
	* •			No. 1
	¥.,			, ·
	·		•	

GCC ACA CTG TGA GTTGGTCAGG GGGGGCTTAC TCGGCGTTTT CCGACACTGC Ala Thr Leu 589	1835
GTTGGTTGCG GCAGTGCGCA CCCCCTGGAT TGATTGCGGG GGTGCCCTGT CGCTGGTGTC	1895
GCCTATCGAC TTAGGGGTAA AGGTCGCTCG CGAAGTTCTG ATGCGTGCGT CGCTTGAACC	1955
ACAAATGGTC GATAGCGTAC TCGCAGGCTC TATGGCTCAA GCAAGCTTTG ATGCTTACCT	2015
GCTCCCGCGG CACATTGGCT TGTACAGCGG TGTTCCCAAG TCGGTTCCGG CCTTGGGGGT	2075
GCAGCGCATT TGCGGCACAG GCTTCGAACT GCTTCGGCAG GCCGGCGAGC AGATTTCCCA	2135
AGGCGCTGAT CACGTGCTGT GTGTCGCGGG CTGCAG	2171
FIG. 2h:	

				-
				•
		1		
7 :				
	•			Ž.
				•
			•	

CTGCAGCCGA (	GCATCGATTG A	GCACTTTAC CC	CAGCTGCGC	TGGCTGACCA TTCAGAATGO	60
CCCGCGGCAC :	TATCCAATCT A	AATCGATCT TO	ceecccc	CGGGCATCAT GCCCGCGGCC	120
CTCGCCTCAT	ITCAATCTCT A	ACTTGATAA AA	AACAGAGCT	GTTCTCCGGT CTTGGTGGAT	т 180
CAAGGCCAGT (	CGCGGAGAGT C	TCGAAGAGG AG	GAGTACAGT	GAACGCCGAG TCCACATTGC	240
AACCGCAGGC	ATCATCATGC I	CTGCTCAGC CA	ACGCTACCG	CAGTGTGTCG ATTGGTCATC	300
CTCCGGTTGA (	GGTTACGCAA G	ACGCTGGAG GT	PATTGTCCG	G ATG CGT TCT CTC GAG Met Arg Ser Leu Glu 1	
				CGT CTC GAG CAT TGG Arg Leu Glu His Trp 20	404
			s Val Ala	GCC AGG GCG GCA AAT Ala Arg Ala Ala Asn 35	452
	Arg Arg Ile			TTC CAC AAC GTC CGC Phe His Asn Val Arg 50	500
				TCG GCA GAG CGT CCG Ser Ala Glu Arg Pro 65	548
		Asn Asp Leu		CTT CAG CTG GCA TTT Leu Gln Leu Ala Phe 85	596
				GTG TCT CCT GCT TAT Val Ser Pro Ala Tyr 100	644
			s Leu Arg	CAC ATC GTA GGT CTT His Ile Val Gly Leu 115	692
	Gly Leu Val			GCA CCT TTC CAG CGC Ala Pro Phe Gln Arg 130 133	740
			s Val Asp	GCG CTG TAT CGT GGT Ala Leu Tyr Arg Gly 575	788
	Ser Met Le	G CGT GAC GAG J Arg Asp Glo 585		CTG TGA GTTGGTCAGG Leu 589	837
GGGGGCTTAC	TCGGCGTTTT	CCGACACTGC G	TTGGTTGCG	GCAGTGCGCA CCCCCTGGA	т 897
TGATTGCGGG	GGTGCCCTGT	CGCTGGTGTC G	CCTATCGAC	TTAGGGGTAA AGGTCGCTC	G 957

			·
			1.0
¥)			
			·
*			4
	Ċ		
		*	
9.			
	Cat y		
	*		
			÷
	į.		PW/
			•
	4,	**************************************	
			(4)

WO 00/26355 PCT/EP99/07952

CGAAGTTCTG	ATGCGTGCGT	CGCTTGAACC	ACAAATGGTC	GATAGCGTAC	TCGCAGGCTC	1017
TATGGCTCAA	GCAAGCTTTG	ATGCTTACCT	GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	1077
TGTTCCCAAG	TCGGTTCCGG	CCTTGGGGGT	GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	1137
GCTTCGGCAG	GCCGGCGAGC	AGATTTCCCA	AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGG	1197
CTGCAG						1203

FIG. 2i:

	••	
		_
		•
		٠
		, . ·
		•
7.		
	-	

GAATT	CCC(	CT G	GCGA	CGAA	A GG	GCGG	CAGO	CCG	CATO	GCC	ACGG	CTGC	GGC (	GTAA	CTGAT	60
GCTTG	CGT	A A	TCGT	TAAC	C GT	TTGA	LTAA	CCI	TGCC	AAA	тттс	GGCC	GAG A	AGAAT	CATGC	120
GGGTA	CGC	ст т	TCCG	TGCG	с тт	TGAT	CTGC	GC1	TCC	STGC	СТТС	SAATO	CAG A	AAA.	TAGTT	180
<b>AAT</b> TG	ACA	GA A	CTAT	AGGT	T CG	CAGT	AGCI	TTI	rgcto	CACC	CACC	CAAA	rcc ,	ACAGO	CACTGG	240
GGTGC	ACG		Asn									· Val			AAG Lys	290
GTT G Val G 15																338
AAC G Asn A																386
GTG C																434
GGC G																482
ACC C													CGGG	GACA	AGC	531
AAGC	GAAC	CG G	FAATI	GCCF	AG C	rggg	GCGC	CTO	CTGGT	raag	GTT	GGAA	AGC (	CCTGC	CAAAGT	591
AAAC	rgga	TG (	CTTI	CTTC	SC CO	SCCA.	AGGA:	г ст	GATG	GCGC	AGGG	GATO	CAA (	GATCT	rgatca	651
AGAG!	ACAG	GA 1	rgago	SATCO	ST TI	rcgc								CAC His		703
GGT S																751
CAA ( Gln (																799
CAG (																847
AAT A								Λrg					Ala	ACG Thr		895

			•
	<b>.</b>		÷

		CCT Pro														9	943
		CTG Leu														g	991
		GCT Ala														10	39
		CAT His														10	87
		CGC Arg 140														11	135
		GAT Asp														11	183
		GCC Ala										_				12	231
		CAT His														12	279
		TCT Ser														13	327
		GAC Asp 220											_			13	375
		TGG Trp														14	123
	Ser	CAG Gln				Phe									TGA	14	471
GCG	GGAC	тст	GGGG	TTCG	AA A	TGAC	CGAC	C AA	GCGA	CGCC	(			GGC . Gly		15	525
		TTC Phe					Ser									15	573

		•
	÷	
		•

AAG CGC TGA Lys Arg 275 276	TAAATGCGCC	GGGCCCTCG	CTGCGCCCCC	GGCCTTCCAF	TAATGACAAT	1632
AATGAGGAGT	GCCCAATGTT	TCACGTGCCC	CTGCTTATTG	GTGGTAAGCC	TTGTTCAGCA	1692
TCTGATGAGC	GCACCTTCGA	GCGTCGTAGC	CCGCTGACCG	GAGAAGTGGT	ATCGCGCGTC	1752
GCTGCTGCCA	GTTTGGAAGA	TGCGGACGCC	GCAGTGGCCG	CTGCACAGGC	TGCGTTTCCT	1,812
GAATGGGCGG	CGCTTGCTCC	GAGCGAACGC	CGTGCCCGAC	TGCTGCGAGC	GGCGGATCTT	1872
CTAGAGGACC	GTTCTTCCGA	GTTCACCGCC	GCAGCGAGTG	AAACTGGCGC	AGCGGGAAAC	1932
TGGTATGGGT	TTAACGTTTA	CCTGGCGGCG	GGCATGTTGC	GGGGAATTC		1981
FIG. 2j:						

		4 . ·
		*
		ţ
		•
4. *		•
	÷	
		i.
1,		•
	-	- J

GAATTCCCCT GGCGACGAAA GGGCGGCAGG CCGCATGGCC ACGGCTGGGC GGTAACTGAT	60
GCTTGCGTTA ATCGTTAACC GTTTGAAATT CCTTGCCAAA TTTCGGCGAG AGAATCATGC	120
GGGTACGCCT TTCCGTGCGC TTTGATCTGC GCTTCCGTGC CTTGAATCAG AAAAATAGTT	180
AATTGACAGA ACTATAGGTT CGCAGTAGCT TTTGCTCACC CACCAAATCC ACAGCACTGG	240
GGTGCACG ATG AAT AGC TAC GAT GGC CGT TGG TCT ACC GTT GAT GTG AAG Met Asn Ser Tyr Asp Gly Arg Trp Ser Thr Val Asp Val Lys  1 5 10	290
GTT GAA GAA GGT ATC GCT TGG GTC ACG CTG AAC CGC CCG GAG AAG CGC Val Glu Glu Gly Ile Ala Trp Val Thr Leu Asn Arg Pro Glu Lys Arg 15 20 25 30	338
AAC GCA ATG AGC CCA ACT CTC AAT CGA GAG ATG GTC GAG GTT CTG GAG Asn Ala Met Ser Pro Thr Leu Asn Arg Glu Met Val Glu Val Leu Glu 35 40 45	386
GTG CTG GAG CAG GAC GCA GAT GCT CGC GTG CTT GTT CTG ACT GGT GCA Val Leu Glu Gln Asp Ala Asp Ala Arg Val Leu Val Leu Thr Gly Ala 50 55 60	434
GGC GAA TCC TGG ACC GCG GGC ATG GAC CTG AAG GAG TAT TTC CGC GAG Gly Glu Ser Trp Thr Ala Gly Met Asp Leu Lys Glu Tyr Phe Arg Glu 65 70 75	482
ACC GAT GCT GGC CCC GAA ATT CTG CAA GAG AAG ATT CGT CGGGGGAGAG Thr Asp Ala Gly Pro Glu Ile Leu Gln Glu Lys Ile Arg 80 85 90 91	531
GCGGTTTGCG TATTGGGCGC ATGCATAAAA ACTGTTGTAA TTCATTAAGC ATTCTGCCGA	591
CATGGAAGCC ATCACAAACG GCATGATGAA CCTGAATCGC CAGCGGCATC AGCACCTTGT	651
CGCCTTGCGT ATAATATTTG CCCATGGACG CACACCGTGG AAACGGATGA AGGCACGAAC	711
CCAGTTGACA TAAGCCTGTT CGGTTCGTAA ACTGTAATGC AAGTAGCGTA TGCGCTCACG	771
CAACTGGTCC AGAACCTTGA CCGAACGCAG CGGTGGTAAC GGCGCAGTGG CGGTTTTCAT	831
GGCTTGTTAT GACTGTTTTT TTGTACAGTC TATGCCTCGG GCATCCAAGC AGCAAGCGCG	891
TTACGCCGTG GGTCGATGTT TGATGTTATG GAGCAGCAAC G ATG TTA CGC AGC AGC Met Leu Arg Ser Ser 1	947
AAC GAT GTT ACG CAG CAG GGC AGT CGC CCT AAA ACA AAG TTA GGT GGC Asn Asp Val Thr Gln Gln Gly Ser Arg Pro Lys Thr Lys Leu Gly Gly 10 15 20	995
TCA AGT ATG GGC ATC ATT CGC ACA TGT AGG CTC GGC CCT GAC CAA GTC Ser Ser Met Gly Ile Ile Arg Thr Cys Arg Leu Gly Pro Asp Gln Val	1043

					) / _ ;
		÷			
÷				į.	
			-	• • •	
				*	

											CGT Arg					1091
											GAT Asp 65					1139
											GCC Ala					1187
											CCC Pro					1235
											GTC Val					1283
											CTC Leu					1331
											CAA Gln 145					1379
											GGC Gly					1427
								AGT Ser			ACC Thr 177	TAA	CAAT	rtcgi	TTC	1476
AAG	CCGA	GAT (	cgc'	rtcc	G.					ys G	AG TT In Pi 50					1526
								ACC Thr		Lys	CGC Arg 276	TGA	TAA	ATGC	GCC	1575
GGG	GCCC	rcg (	CTGC	GCCC	CC G	GCCT'	rccai	A TA	ATGAG	CAAT	AAT	SAGG	AGT (	GCCC/	<b>AATGTT</b>	1635
TCA	CGTG	ccc (	CTGC'	TAT	TG G	rggt	AAGC	C TT	GTTC/	AGCA	TCT	SATG	AGC (	GCAC	CTTCGA	1695
GCG	TCGT	AGC (	CCGC'	TGAC	CG G	AGAA	GTGG'	T AT	CGCGG	CGTC	GCT	CTG	CCA (	STTTC	GGAAGA	1755
TGC	GGAC	GCC (	GCAG	TGGC	CG C	rgca	CAGG	C TG	CGTT	CCT	GAA:	rggg	CGG (	CGCT	rgctcc	1815
GAG	CGAA	CGC (	CGTG	CCCG.	AC T	GCTG	CGAG	C GG	CGGA:	rctt	CTA	GAGG	ACC (	GTTC:	TTCCGA	1875

					,
				-	
			4.		
		~			
•		· ·			

GTTCACCGCC GCAGCGAGTG AAACTGGCGC AGCGGGAAAC TGGTATGGGT TTAACGTTTA 1935
CCTGGCGGCG GGCATGTTGC GGGGAATTC 1964

29

PCT/EP99/07952

FIG. 2k:

WO 00/26355

••			•)		
					-
					ν.
ÿ			i.		
		. T		~	
		1.		•	

GAATTCCCCT GGCGACGAAA GGGCGGCAGG CCGCATGGCC ACGGCTGGGC GGTAACTGAT	60
GCTTGCGTTA ATCGTTAACC GTTTGAAATT CCTTGCCAAA TTTCGGCGAG AGAATCATGC	120
GGGTACGCCT TTCCGTGCGC TTTGATCTGC GCTTCCGTGC CTTGAATCAG AAAAATAGTT	180
AATTGACAGA ACTATAGGTT CGCAGTAGCT TTTGCTCACC CACCAAATCC ACAGCACTGG	240
GGTGCACG ATG AAT AGC TAC GAT GGC CGT TGG TCT ACC GTT GAT GTG AAG  Met Asn Ser Tyr Asp Gly Arg Trp Ser Thr Val Asp Val Lys  1 5 10	290
GTT GAA GAA GGT ATC GCT TGG GTC ACG CTG AAC CGC CCG GAG AAG CGC Val Glu Glu Gly Ile Ala Trp Val Thr Leu Asn Arg Pro Glu Lys Arg 15 20 25 30	338
AAC GCA ATG AGC CCA ACT CTC AAT CGA GAG ATG GTC GAG GTT CTG GAG Asn Ala Met Ser Pro Thr Leu Asn Arg Glu Met Val Glu Val Leu Glu 35 40 45	386
GTG CTG GAG CAG GAC GCA GAT GCT CGC GTG CTT GTT CTG ACT GGT GCA  Val Leu Glu Gln Asp Ala Asp Ala Arg Val Leu Val Leu Thr Gly Ala  50 55 60	434
GGC GAA TCC TGG ACC GCG GGC ATG GAC CTG AAG GAG TAT TTC CGC GAG Gly Glu Ser Trp Thr Ala Gly Met Asp Leu Lys Glu Tyr Phe Arg Glu 65 70 75	482
ACC GAT GCT GGC CCC GAA ATT CTG CAA GAG AAG ATT CGT CGC GAG CAG Thr Asp Ala Gly Pro Glu Ile Leu Gln Glu Lys Ile Arg Arg Glu Gln 80 85 90 92 255	530
GGC ATG AAG CAG TTC CTT GAC GAG AAA AGC ATC AAG CCG GGC TTG CAG Gly Met Lys Gln Phe Leu Asp Glu Lys Ser Ile Lys Pro Gly Leu Gln 260 265 270	578
ACC TAC AAG CGC TGA TAAATGCGCC GGGGCCCTCG CTGCGCCCCC GGCCTTCCAA Thr Tyr Lys Arg 275 276	633
TAATGACAAT AATGAGGAGT GCCCAATGTT TCACGTGCCC CTGCTTATTG GTGGTAAGCC	693
TTGTTCAGCA TCTGATGAGC GCACCTTCGA GCGTCGTAGC CCGCTGACCG GAGAAGTGGT	753
ATCGCGCGTC GCTGCTGCCA GTTTGGAAGA TGCGGACGCC GCAGTGGCCG CTGCACAGGC	813
TGCGTTTCCT GAATGGGCGG CGCTTGCTCC GAGCGAACGC CGTGCCCGAC TGCTGCGAGC	873
GGCGGATCTT CTAGAGGACC GTTCTTCCGA GTTCACCGCC GCAGCGAGTG AAACTGGCGC	933
AGCGGGAAAC TGGTATGGGT TTAACGTTTA CCTGGCGGCG GGCATGTTGC GGGGAATTC FIG. 21:	992

		· S		7	
					er <sub>E</sub>
		ů.			
	2				
				÷.	

GAATTCCAAT AATGACAATA ATGAGGAGTG CCCA ATG TTT CAC GTG CCC CTG CTT  Met Phe His Val Pro Leu Leu  1 5												55			
	GGT Gly														103
	AGC Ser 25														151
	GAA Glu														199
	TGG Trp														247
	GCG Ala														295
	GAA Glu														343
	GCG Ala 105														391
	GAT Asp														439
	CAG Gln														487
	ATC Ile			Val											535
	GTG Val		Leu					Leu					His		583
	GGT Gly 185	Gln					Ala					Gly			631
	ATC Ile					Gln					Val				679

ķ.		
	Δ <sub>1</sub>	
	: 1 :	
		•
	*	
		•

ATT GCA AAT CCT GCG GTA CGT CGA GTG A Ile Ala Asn Pro Ala Val Arg Arg Val A 220 . 2	_	7
GTT GGA CGG ATC ATT GGT GAG CTG TCT G Val Gly Arg Ile Ile Gly Glu Leu Ser A 235		5
GTG CTG GAA TTA GGT GGT AAG GCT CCG T Val Leu Glu Leu Gly Gly Lys Ala Pro F 250 255		3
GAC CTC GAT GCG GCG GTC GAA GCG GCG GASP Leu Asp Ala Ala Val Glu Ala Ala Ala Ala 265 270		1
CAG GGT CAA ATC TGC ATG TCC ACT GAG CGIn Gly Gln Ile Cys Met Ser Thr Glu A		9
GCA GAC GCC TTT GTT GAA AAG CTG GCG A Ala Asp Ala Phe Val Glu Lys Leu Ala A 300		.7
GCT GGC GAT CCT AAT GAT CCG CAA TCG G Ala Gly Asp Pro Asn Asp Pro Gln Ser V 315		5
GCC AAT GCA GGT CAA CGC ATC CAG GTT C Ala Asn Ala Gly Gln Arg Ile Gln Val I 330 335		3
GACAGCAAGC GAACCGGAAT TGCCAGCTGG GGCC	GCCCTCT GGTAAGGTTG GGAAGCCCTG 112	:3
CAAAGTAAAC TGGATGGCTT TCTTGCCGCC AAGC	GATCTGA TGGCGCAGGG GATCAAGATC 118	3
TGATCAAGAG ACAGGATGAG GATCGTTTCG C AT	TG ATT GAA CAA GAT GGA TTG 123 et Ile Glu Gln Asp Gly Leu 1 5	5
CAC GCA GGT TCT CCG GCC GCT TGG GTG GH is Ala Gly Ser Pro Ala Ala Trp Val G		3
TGG GCA CAA CAG ACA ATC GGC TGC TCT ( Trp Ala Gln Gln Thr Ile Gly Cys Ser A 25 30		11
TCA GCG CAG GGG CGC CCG GTT CTT TTT C Ser Ala Gln Gly Arg Pro Val Leu Phe V 40 45		19
GCC CTG AAT GAA CTG CAG GAC GAG GCA GAA Ala Leu Asn Glu Leu Gln Asp Glu Ala 60		?7

		As-			٠	,
· ·						
ÿ						
		ŕ				
				~		
	¥				·	

												ACT Thr 85			1475
												GAT Asp			1523
												GCT Ala			1571
												TTC Phe			1619
												GAA Glu			1667
	 -	-										CTC Leu 165			1715
	 											GGC Gly			1763
												ATG Met			1811
												GGT Gly			1859
												GCT Ala			1907
								Leu				GGT Gly 245			1955
_	 	Ser					Phe					GAC Asp			2003
TTC Phe 264	GCG	GGAC	TCT	GGGG	TTCG	AA A	TGAC	CGAC	C AA	GCGA	CGCC		GCC Ala 421		2057
		Ser					lle					Val		GAC Asp	2105

			* 4		
		•			
	`				
· ·			-	•	

GAG GCT CAG ATG CCA TTC GGT GGG GTG AAG TCC AGC GGC TAC GGC AGG Glu Ala Gln Met Pro Phe Gly Gly Val Lys Ser Ser Gly Tyr Gly Se 440 445 450	
TTC GGC AGT CGA GCA TCG ATT GAG CAC TTT ACC CAG CTG CGC TGG CT Phe Gly Ser Arg Ala Ser Ile Glu His Phe Thr Gln Leu Arg Trp Le 455 460 465 47	
ACC ATT CAG AAT GGC CCG CGG CAC TAT CCA ATC TAA ATCGATCTTC Thr Ile Gln Asn Gly Pro Arg His Tyr Pro Ile 475 480 481	2247
GGGCGCCGCG GGCATCATGC CCGCGGCGCT CGCCTCATTT CAATCTCTAA CTTGATA	AAAA 2307
ACAGAGCTGT TCTCCGGTCT TGGTGGATCA AGGCCAGTCG CGGAGAGTCT CGAAGAG	GGAG 2367
AGTACAGTGA ACGCCGAGTC CACATTGCAA CCGCAGGCAT CATCATGCTC TGCTCAC	GCCA 2427
CGCTACCGCA GTGTGTCGAT TGGTCATCCT CCGGTTGAGG TTACGCAAGA CGCTGGA	AGGT 2487
ATTGTCCGGA TGCGTTCTCT CGAGGCGCTT CTTCCCTTCC	2539
FIG. 2m:	

	÷			
		N <sub>i</sub>		
			•	

GAATTCCAAT AATGACAATA ATGAGGAGTG CCCA ATG TTT CAC GTG CCC CTG CTT  Met Phe His Val Pro Leu Leu  1 5													55			
	GGT Gly														1	103
	AGC Ser 25														1	151
	GAA Glu														3	199
	TGG Trp														2	247
	GCG Ala														2	295
	GAA Glu														:	343
	GCG Ala 105														:	391
	GAT Asp															439
	CAG Gln				Val											487
	ATC Ile			Val					Met							535
	GTG Val		Leu					Leu					His			583
	GGT Gly 185	Gln					Ala					Gly				631
	C ATC L Ile					Gln					Val					679

		-

												GGT Gly				727
												CTG Leu		Pro		775
												TTG Leu 260				823
												GCC Ala				871
												GTG Val				919
												GCC Ala				967
												TCG Ser				1015
			GGT Gly					GTG	GGGA	GAG (	GCGG.	TTTG	CG T	ATTGO	GCGC	1069
ATG	САТА	AAA .	ACTG	TTGT	AA T	TCAT	TAAG	C AT	TCTG	CCGA	CAT	GGAA	GCC /	ATCAG	CAAACG	1129
GCA	TGAT	GAA	CCTG	AATC	GC C	AGCG	GCAT	C AG	CACC'	TTGT	CGC	CTTG	CGT .	ATAA	TATTTG	1189
ccc	ATGG	ACG	CACA	CCGT	GG A	AACG	GATG	A AG	GCAC	GAAC	CCA	GTTG.	ACA	TAAG	CCTGTT	1249
CGG	TTCG	TAA	ACTG	TAAT	GC A	AGTA	GCGT	A TG	CGCT	CACG	CAA	CTGG	TCC .	AGAA	CCTTGA	1309
CCG	AACG	CAG	CGGT	GGTA	AC G	GCGC	agtg	G CG	GTTT	TCAT	GGC	TTGT	TAT	GACT	GTTTTT	1369
TTG	TACA	GTC	TATG	сстс	GG G	CATC	CAAG	C AG	CAAG	CGCG	TTA	CGCC	GTG	GGTC	GATGTT	1429
TGA	TGTT	ATG	GAGC	AGCA	AC G		Leu				Asn	GAT Asp				1480
					Lys					Gly		AGT Ser				1528
				Arg					Gln			TCC Ser		Arg		1576

			• • • • • • • • • • • • • • • • • • •
		÷	
	4.		
		,	
**		•	

GCT CTT GAT CTT TTC GGT CGT GAG TTC GGA GAC GTA GCC ACC TAC TCC Ala Leu Asp Leu Phe Gly Arg Glu Phe Gly Asp Val Ala Thr Tyr Ser 45 50 55	1624
CAA CAT CAG CCG GAC TCC GAT TAC CTC GGG AAC TTG CTC CGT AGT AAG Gln His Gln Pro Asp Ser Asp Tyr Leu Gly Asn Leu Leu Arg Ser Lys 60 65 70	1672
ACA TTC ATC GCG CTT GCT GCC TTC GAC CAA GAA GCG GTT GTT GGC GCT Thr Phe Ile Ala Leu Ala Ala Phe Asp Gln Glu Ala Val Val Gly Ala 75 80 85 90	1,720
CTC GCG GCT TAC GTT CTG CCC AGG TTT GAG CAG CCG CGT AGT GAG ATC Leu Ala Ala Tyr Val Leu Pro Arg Phe Glu Gln Pro Arg Ser Glu Ile 95 100 105	1768
TAT ATC TAT GAT CTC GCA GTC TCC GGC GAG CAC CGG AGG CAG GGC ATT  Tyr Ile Tyr Asp Leu Ala Val Ser Gly Glu His Arg Arg Gln Gly Ile  110 115 120	1816
GCC ACC GCG CTC ATC AAT CTC CTC AAG CAT GAG GCC AAC GCG CTT GGT Ala Thr Ala Leu Ile Asn Leu Lys His Glu Ala Asn Ala Leu Gly 125 130 135	1864
GCT TAT GTG ATC TAC GTG CAA GCA GAT TAC GGT GAC GAT CCC GCA GTG Ala Tyr Val Ile Tyr Val Gln Ala Asp Tyr Gly Asp Asp Pro Ala Val 140 145	1912
GCT CTC TAT ACA AAG TTG GGC ATA CGG GAA GAA GTG ATG CAC TTT GAT Ala Leu Tyr Thr Lys Leu Gly Ile Arg Glu Glu Val Met His Phe Asp 155 160 165 170	1960
ATC GAC CCA AGT ACC GCC ACC TAA CAATTCGTTC AAGCCGAGAT CGGCTTCCCA Ile Asp Pro Ser Thr Ala Thr 175 177	2014
A TTG GCC CAG CGC GTC GAT TCG GGC ATT TGC CAT ATC AAT GGA CCG ACT Leu Ala Gln Arg Val Asp Ser Gly Ile Cys His Ile Asn Gly Pro Thr 420 425 430 435	2063
GTG CAT GAC GAG GCT CAG ATG CCA TTC GGT GGG GTG AAG TCC AGC GGC Val His Asp Glu Ala Gln Met Pro Phe Gly Gly Val Lys Ser Ser Gly 440 445 450	2111
TAC GGC AGC TTC GGC AGT CGA GCA TCG ATT GAG CAC TTT ACC CAG CTG Tyr Gly Ser Phe Gly Ser Arg Ala Ser Ile Glu His Phe Thr Gln Leu 455 460 465	2159
CGC TGG CTG ACC ATT CAG AAT GGC CCG CGG CAC TAT CCA ATC TAA Arg Trp Leu Thr Ile Gln Asn Gly Pro Arg His Tyr Pro Ile 470 475 480 481	2204
ATCGATCTTC GGGCGCCGCG GGCATCATGC CCGCGGCGCT CGCCTCATTT CAATCTCTAA	2264
CTTGATAAAA ACAGAGCTGT TCTCCGGTCT TGGTGGATCA AGGCCAGTCG CGGAGAGTCT	2324

			• •
			میر م
.*)		931	

WO 00/26355 PCT/EP99/07952

CGAAGAGGAG	AGTACAGTGA	ACGCCGAGTC	CACATTGCAA	CCGCAGGCAT	CATCATGCTC	2384
TGCTCAGCCA	CGCTACCGCA	GTGTGTCGAT	TGGTCATCCT	CCGGTTGAGG	TTACGCAAGA	2444
CGCTGGAGGT	ATTGTCCGGA	TGCGTTCTCT	CGAGGCGCTT	CTTCCCTTCC	CGGGTGGAAT	2504
TC						2506
FIG. 2n:						

				-
	*			
i.				
,				
		•	• • • • • • • • • • • • • • • • • • • •	
			<b>1</b>	

GAATTCCAAT AATGACAATA ATGAGGAGTG CCCA ATG TTT CAC GTG CCC CTG CTT  Met Phe His Val Pro Leu Leu  1 5										55				
			AAG Lys											103
			CTG Leu											151
			GCG Ala											199
			GCG Ala											247
			CTT Leu 75											295
			GGC Gly											343
			ATG Met											391
_			ATT Ile											439
			TGT Cys											487
			GGC Gly 155											535
			TTG Leu					_						583
		Gln	GTG Val				Ala				Gly		AAT Asn	631
	Ile					Gln				Val			CTG Leu 215	679

			\$.2.	
				-
9.				
	•*			

ATT GC															727
GTT GG Val Gl															775
GTG CT Val Le															823
GAC CT Asp Le 26			-				_	_							871
CAG GG Gln Gl 280	GT CAA ly Gln														919
	AC GCC sp Ala														967
	GC GAT ly Asp														1015
	AT GCA sn Ala 330	Gly													1063
Lys G	GC GCG ly Ala 45 346		TGGA		u Al					Se:				C CAT s His 430	1113
	AT GGA sn Gly			Val											1161
	AG TCC ys Ser		Gly					Gly							1209
	TT ACC he Thr 465	Gln					Thr								1257
Tyr P	CA ATC ro Ile 80 481	:	ATC	GATC	TTC	GGGC	GCCG	CG G	GCAT	CATG	c cc	GCGG	CGCT		1309
CGCCT	CATTT	CAAT	стст	'AA C	TTGA	TAAA	A AC	AGAG	CTGT	TCT	ccgg	TCT	TGGT	GGATCA	1369
AGGCC	AGTCG	CGGA	GAGI	CT C	GAAG	AGGA	G AG	TACA	GTGA	ACG	CCGA	GTC	CACA	TTGCAA	1429

			9	
				-
		*		
			* 4.4	
· ·				
	i.			
			<b>,</b>	
			<b>~</b>	

WO 00/26355 PCT/EP99/07952

CCGCAGGCAT	CATCATGCTC	TGCTCAGCCA	CGCTACCGCA	GTGTGTCGAT	TGGTCATCCT	1489
CCGGTTGAGG	TTACGCAAGA	CGCTGGAGGT	ATTGTCCGGA	TGCGTTCTCT	CGAGGCGCTT	1549
CTTCCCTTCC	CGGGTGGAAT	TC			÷	1571
FIG. 20:						

					e.	
3						
,						
	2					
				•		

GAAT	TCCG	CG (	STCGG	CGA	AA GI	TGAT	rgcgc	TGI	CATC	STGG	TGA	AGAT	CAA '	TCCA'	rgci	GC.	60
GTGA	CGAC	GC (	CACAC					CA GC er Gl					er A				112
			TTG Leu														160
			TCG Ser														208
			CTG Leu													•	256
			GGC Gly														304
			ATT Ile 80														352
			CAG Gln														400
			CAG Gln														448
			ATG Met													,	496
			CTC Leu														544
			GAT Asp 160	Pro										Ala			592
	CTG Leu 174		ACAG(	CAA	GCGA	ACCG	GA A'	TTGC	CAGC	T GG	GGCG	СССТ	CTG	GTAA	GGT		648
TGG	GAAG	ccc	TGCA	aagt	AA A	CTGG.	ATGG	C TT	TCTT	GCCG	CCA	AGGA	тст	GATG	GCG	CAG	708
GGG	ATCA	AGA	TCTG	ATCA	AG A	GACA	GGAT	G AG	GATC	GTTT	CGC			GAA Glu		_	763

			÷
		4	
HE:			
	*		
		•	
	•		

			CAC His													811
			TGG Trp													859
			TCA Ser 40													.907
			GCC Ala													955
			ACG Thr													1003
Thr 85	Glu	Ala	GGA Gly	Arg	Asp 90	Trp	Leu	Leu	Leu	Gly 95	Glu	Val	Pro	Gly	Gln 100	1051
Asp	Leu	Leu	TCA Ser	Ser 105	His	Leu	Ala	Pro	Ala 110	Glu	Lys	Val	Ser	Ile 115	Met	1099
			ATG Met 120													1147
Phe	Asp	His 135	CAA Gln	Ala	Lys	His	Arg 140	Ile	Glu	Arg	Ala	Arg 145	Thr	Arg	Met	1195
Glu	Ala 150	Gly	CTT Leu	Val	Asp	Gln 155	Asp	Asp	Leu	Asp	Glu 160	Glu	His	Gln	Gly	1243
Leu 165	Ala	Pro	GCC Ala	Glu	Leu 170	Phe	Ala	Arg	Leu	Lys 175	Ala	Arg	Met	Pro	Asp 180	1291
			CTC Leu		Val					Ala					Ile	1339
Met	Val	Glu	AAT Asn 200	Gly	Arg	Phe	Ser	Gly 205	Phe	Ile	Asp	Cys	Gly 210	Arg	Leu	1387
			GAC Asp					lle					Arg			1435

	•	
	-	 ÷

					Gly											1483
					GAT Asp 250											1531
		TTC Phe		TGA	GCGG	GACT	CT G	GGGT	TCGA	TA AA	GACC	CGACC	CAA C	GCGAC	CGCC	1586
I			Ala G		SAG C			rp I					le V			1633
		-			TTC Phe											1681
					GCA Ala											1729
					GCC Ala											1777
					GGC Gly 265											1825
					CGA Arg											1873
					TCC Ser											1921
					GCG Ala											1969
	_	Arg			GAC Asp											2017
	Val				CAG Gln 345											2065
					GCC					His					Thr	2113

			1.2	
			,	
			(a	
		~		

GGA Gly	TTG Leu	CGT Arg	CTC Leu 375	TGC Cys	ATG Met	ACC Thr	CTC Leu	GCT Ala 380	CAC His	CAA Gln	TTG Leu	CAA Gln	GCT Ala 385	AAT Asn	AAC Asn	2161
TTT Phe	CGA Arg	TAT Tyr 390	GGA Gly	ATT Ile	GCC Ala	TCG Ser	GCA Ala 395	TGC Cys	ATT Ile	GGT Gly	GGG Gly	GGA Gly 400	CAG Gln	GGG Gly	ATG Met	2209
GCG Ala	GTT Val 405	CTT Leu	TTA Leu	GAG Glu	AAT Asn	CCC Pro 410	CAC His	TTC Phe	GGT Gly	TCG Ser	TCC Ser 415	TCT Ser	GCA Ala	CGA Arg	AGT Ser	2257
					GTT Val 425	Asp				Leu		TAA	CGG	GCAT	CTC	2306
CTT	TGTT	GCT	TTGA	GGTG	GC G	CACG	AAGG	A GG	GCTC	GAAA	ATC'	rctg	CTA .	AAAA -	CAAGAA	2366
GAA	GGAA	CAG	GGAA	CATG	AT T	AGTT	TCGC	T CG	TATG	GCAG	AAA	GTTT.	AGG	AGTC	CAGGCT	2426
AAA	CTTG	ccc	TTGC	CTTC	GC A	CTCG	TATT	A TG	TGTC	GGGC	TGA	TTGT	TAC	CGGC.	ACGGGT	2486
TTC	TACA	GTG	TACA	TACC	TT G	TCAG	GGTT	G GT	GGGA	ATTC						2526
FIG	. 2p	:														

				<i>i.</i> .	
					1.0
			10. 7		
A					

GAAT	TCCG	CG C	TCGG	CGA	A GI	TGAT	GCGC	TGT	'ATC	TGG	TGAA	AGATO	CAA 1	CCAI	CGCTGC	60
GTGA	CGAG	GC C	CACAC										er Al		T TCC ne Ser	112
-												ATT Ile 25				160
-		-										GTA Val				208
												ATG Met				256
												GCT Ala				304
												TCG Ser				352
												CTG Leu 105				400
												CTG Leu				448
												ACA Thr				496
												TTT Phe				544
GCA Ala				Pro								GCT Ala				592
	CTG Leu 174		GAGA	GGC	GGTT'	TGCGʻ	ГА Т	rggg	CGCA'	r GC	АТАА	AAAC	TGT'	TGTA.	ATT	648
CAT	TAAG	CAT	TCTG	CCGA	CA T	GGAA	GCCA'	T CA	CAAA	CGGC	ATG	ATGA	ACC	TGAA'	TCGCCA	708
GCG	GCAT	CAG	CACC	TTGT	CG C	СТТС	CGTA	AA T	TATT	TGCC	CAT	GGAC	GCA	CACC	GTGGAA	768
ACG	GATG	AAG	GCAC	GAAC	CC A	GTTG.	ACAT	A AG	CCTG	TTCG	GTT	CGTA	AAC	TGTA	ATGCAA	828
GTA	GCGT	ATG	CGCT	CACG	CA A	CTGG	TCCA	G AA	CCTT	GACC	GAA	CGCA	GCG	GTGG	TAACGG	888

			=\frac{1}{2}.	
				-
			i.	
	e)			
, a 12 - 14 in				
		-		

CGCAGTGGCG GTTTTC	ATGG CTTGTTATG	A CTGTTTTTT	GTACAGTCTA TGCCTCGGGC	948
ATCCAAGC AGCAAGCG	CG TTACGCCGTG	GGTCGATGTTTG	ATGTTATGGA GCAGCAACG	1007
			GGC AGT CGC CCT AAA Gly Ser Arg Pro Lys 15	1055
			CGC ACA TGT AGG CTC Arg Thr Cys Arg Leu 30	1103
		Arg Ala Ala	CTT GAT CTT TTC GGT Leu Asp Leu Phe Gly 45	1151
			CAT CAG CCG GAC TCC His Gln Pro Asp Ser 60	1199
			TTC ATC GCG CTT GCT Phe Ile Ala Leu Ala 80	1247
			GCG GCT TAC GTT CTG Ala Ala Tyr Val Leu 95	1295
			ATC TAT GAT CTC GCA Ile Tyr Asp Leu Ala 110	1343
		n Gly Ile Ala	ACC GCG CTC ATC AAT Thr Ala Leu Ile Asn 125	1391
			TAT GTG ATC TAC GTG Tyr Val Ile Tyr Val 140	1439
			CTC TAT ACA AAG TTG Leu Tyr Thr Lys Leu 160	1487
Gly Ile Arg Glu (			GAC CCA AGT ACC GCC Asp Pro Ser Thr Ala 175	1535
ACC TAA CAATTCGTT Thr 177	TC AAGCCGAGAT		G AGG. GCG CAA GAG GAG u Arg Ala Gln Glu Glu 7 200	
		l Ala Val Thr	GAT GAA CAG TTC GAT Asp Glu Gln Phe Asp 215	1637

		•
		7
	•	

	TAC Tyr												1685
	GTG Val												1733
 	AAG Lys												1,781
	GTA Val 270												1829
	ACA Thr												1877
	GAG Glu												1925
 	CTT Leu												1973
	AAC Asn												2021
	ATT Ile 350												2069
	CAC His												2117
 	CAA Gln												2165
Cys												CCC Pro 410	2213
	TCG Ser						Met					GAC Asp	2261
	CTG Leu 430	TAA	CGG	GCAT	CTC	СТ <b>Т</b> Т	GTTG	СТ Т	TGAG	GTGG	С		2309

			-
		*14.	
			4.
		<b>~</b>	

WO 00/26355 PCT/EP99/07952

GCACGAAGGA	GGGCTCGAAA	ATCTCTGCTA	AAAACAAGAA	GAAGGAACAG	GGAACATGAT	2369
TAGTTTCGCT	CGTATGGCAG	AAAGTTTAGG	AGTCCAGGCT	AAACTTGCCC	TTGCCTTCGC	2429
ACTCGTATTA	TGTGTCGGGC	TGATTGTTAC	CGGCACGGGT	TTCTACAGTG	TACATACCTT	2489
GTCAGGGTTG	GTGGGAATTC					2509
FIG. 2q:						

•		
	\$ <sub>10</sub>	
	-	

GAATTCCGCG GTCGC	SCGAAA GTTGATO	GCGC TGTATCGTGG	TGAAGATCAA TCCA	IGCTGC 60
GTGACGAGGC CACAC			CT TAC TCG GCG Tools Tyr Ser Ala Pl	
			TGG ATT GAT TGC Trp Ile Asp Cys 25	
			GGG GTA AAG GTC Gly Val Lys Val 40	
			CAA ATG GTC GAT Gln Met Val Asp	
			GAT GCT TAC CTG Asp Ala Tyr Leu 75	50,
			AAG TCG GTT CCG Lys Ser Val Pro 90	
	Arg Ile Cys		GAA CTG CTT CGG Glu Leu Leu Arg 105	
			GTG CTG TGT GTC Val Leu Cys Val 120	
			TAT ACA CAC CGG Tyr Thr His Arg	
			GAT TTT TTG TGG Asp Phe Leu Trp 155	
			ATC GCT ACC GCA Ile Ala Thr Ala 170	
AAC CTG GCG CGC Asn Leu Ala Arg 175 176			AA TGG ATT GAC C ys Trp Ile Asp G 205	
			GAG GGC TAC AAC Glu Gly Tyr Asn 220	

``				.º	
			:		
					3
		.*·			

<b>-</b>	
ATC CGC GGC CTA GCA GTC TTT GAA GCC CTT TCC CGA TTG AAG CCT GTT Ile Arg Gly Leu Ala Val Phe Glu Ala Leu Ser Arg Leu Lys Pro Val 245 250 255	785
CAT TCT GGC GGG GTG CAG ACT GCG GGC AAC AGC TGT GCC GTA GTG GAC His Ser Gly Gly Val Gln Thr Ala Gly Asn Ser Cys Ala Val Val Asp 260 265 270	833
GGC GCC GCG GCT TTG GTG GCT CGA GAG TCG TCT GCG ACA CAG CCG Gly Ala Ala Ala Leu Val Ala Arg Glu Ser Ser Ala Thr Gln Pro 275 280 285	881
GTC TTG GCT AGG ATA CTG GCT ACC TCC GTA GTC GGG ATC GAG CCC GAG Val Leu Ala Arg Ile Leu Ala Thr Ser Val Val Gly Ile Glu Pro Glu 290 295 300	. 929
CAT ATG GGG CTC GGC CCT GCG CCC GCG ATT CGC CTG CTT GCG CGT His Met Gly Leu Gly Pro Ala Pro Ala Ile Arg Leu Leu Ala Arg 305 310 315 320	977
AGT GAT CTT AGT TTG AGG GAT ATC GAC CTC TTT GAG ATA AAC GAG GCG Ser Asp Leu Ser Leu Arg Asp Ile Asp Leu Phe Glu Ile Asn Glu Ala 325 330 335	1025
CAG GCC GCC CAA GTT CTA GCG GTA CAG CAT GAA TTG GGT ATT GAG CAC Gln Ala Ala Gln Val Leu Ala Val Gln His Glu Leu Gly Ile Glu His 340 345 350	1073
TCA AAA CTT AAT ATT TGG GGC GGG GCC ATT GCA CTT GGA CAC CCG CTT Ser Lys Leu Asn Ile Trp Gly Gly Ala Ile Ala Leu Gly His Pro Leu 355 360 365	1121
GCC GCG ACC GGA TTG CGT CTC TGC ATG ACC CTC GCT CAC CAA TTG CAA Ala Ala Thr Gly Leu Arg Leu Cys Met Thr Leu Ala His Gln Leu Gln 370 375 380	1169
GCT AAT AAC TTT CGA TAT GGA ATT GCC TCG GCA TGC ATT GGT GGG GGA Ala Asn Asn Phe Arg Tyr Gly Ile Ala Ser Ala Cys Ile Gly Gly 385 390 395 400	1217
CAG GGG ATG GCG GTT CTT TTA GAG AAT CCC CAC TTC GGT TCG TCC TCT Gln Gly Met Ala Val Leu Leu Glu Asn Pro His Phe Gly Ser Ser Ser 405 410 415	1265
GCA CGA AGT TCG ATG ATT AAC AGA GTT GAC CAC TAT CCA CTG AGC TAA Ala Arg Ser Ser Met Ile Asn Arg Val Asp His Tyr Pro Leu Ser 420 425 430 431	1313
CGGGCATCTC CTTTGTTGCT TTGAGGTGGC GCACGAAGGA GGGCTCGAAA ATCTCTGCTA	1373
AAAACAAGAA GAAGGAACAG GGAACATGAT TAGTTTCGCT CGTATGGCAG AAAGTTTAGG	1433
AGTCCAGGCT AAACTTGCCC TTGCCTTCGC ACTCGTATTA TGTGTCGGGC TGATTGTTAC	1493
CGGCACGGGT TTCTACAGTG TACATACCTT GTCAGGGTTG GTGGGAATTC	1543

FIG. 2r:

		•	
		, j	
	5		
		₹.	
		· .	

			_			
	GGCTGAAAAG					60
	GCTCGATGGC				GGAGAGTTCG	120
	TAAATTTGCT				TGATTTTCTG	180
	CATGAAATTC		CACTTTTCGG		CACGGGATTG	240
	GGAGAGTGCA				TCAGGATGGT	300
GCATGGAAAT	GGCATGAGCT	TTGCTGGATA	TGATTAGAGA		TTTGGCGGAA	360
	ATTCCTCGCC		GTAACCGCGA	CATTCAGGAC	CGTAAAAAGG	420
	CAACTGACCA		CGTCGTCACC	GGAGTGTCCT	CCGGTATCGG	480
	GCCCGCGTTC		CGGCGCCACA	GTGATTGGCG	TAGATCGCAA	540
CATGCCGAGC	CTGACTCTGG	ATGCTTTCGT		CTGAGCCATC		600
	ATCGGGACAG			GCTGGGGCGC		660
	CCCTGCAAAG		GGCTTTCTTG	CCGCCAAGGA	TCTGATGGCG	720
CAGGGGATCA	AGATCTGATC	AAGAGACAGG	ATGAGGATCG	TTTCGCATGA	TTGAACAAGA	780
	GCAGGTTCTC		GGTGGAGAGG	CTATTCGGCT	ATGACTGGGC	840
ACAACAGACA	ATCGGCTGCT	CTGATGCCGC	CGTGTTCCGG	CTGTCAGCGC	AGGGGCGCCC	900
GGTTCTTTTT	GTCAAGACCG	ACCTGTCCGG	TGCCCTGAAT	GAACTGCAGG	ACGAGGCAGC	960
	TGGCTGGCCA			GCTGTGCTCG		1020
TGAAGCGGGA	AGGGACTGGC					1080
TCACCTTGCT			CATGGCTGAT			1140
	GCTACCTGCC					1200
	GAAGCCGGTC					1260
	GAACTGTTCG					1320
CGTGACCCAT	GGCGATGCCT					1380
ATTCATCGAC			GGACCGCTAT		CGTTGGCTAC	1440
	GCTGAAGAGC				TGCTTTACGG	1500
	CCCGATTCGC					1560
	TGGGGTTCGA					1620
	TGAGGAGTCA					1680
	CTACGTGTAA					1740
	AGCAGCTTTG					1800
	AGGATTCTTG					1860
	GAAAGCAAGT					1920
	AGCGCTGCCA					1980
	GGTGCGTTGG					2040
	CAAGTCTGAC					2100
GTCAGGGTGA	TCGTAACTTT	GACCGGGGGC	TTGGTATCCA	ATCGTCTCGA	TATTCTGGCT	2160
GCAG						2164

÷	
	-

				GAAGGGAGGG		60
				GCGCGGTCTT		120
				ATGGGTTGGA		180
CATTCTGCAT				GGGGTGGGTG	CACGGGATTG	240
AAGGTTGCTA			GCCCAGGAAG	CACGCGGGTT	TCAGGATGGT	300
GCATGGAAAT		TTGCTGGATA		CATTAACTAT	TTTGGCGGAA	360
TGGAAGCACG	ATTCCTCGCC				CGTAAAAAGG	420
AAAGAGCATG	CAACTGACCA	ACAAGAAAAT	CGTCGTCACC	GGAGTGTCCT	CCGGTATCGG	480
TGCCGAAACT			CGGCGCCACA		TAGATCGCAA	540
CATGCCGAGC	CTGACTCTGG	ATGCTTTCGT	TCAGGCTGAC	CTGAGCCATC	CTGAGGGGAG	600
AGGCGGTTTG	CGTATTGGGC	GCATGCATAA	AAACTGTTGT	AATTCATTAA	GCATTCTGCC	660
GACATGGAAG	CCATCACAAA	CGGCATGATG	AACCTGAATC	GCCAGCGGCA	TCAGCACCTT	720
GTCGCCTTGC	GTATAATATT	TGCCCATGGA	CGCACACCGT	GGAAACGGAT	GAAGGCACGA	780
ACCCAGTTGA	CATAAGCCTG	TTCGGTTCGT	AAACTGTAAT	GCAAGTAGCG	TATGCGCTCA	840
CGCAACTGGT	CCAGAACCTT	GACCGAACGC	AGCGGTGGTA	ACGGCGCAGT	GGCGGTTTTC	900
ATGGCTTGTT	ATGACTGTTT	TTTTGTACAG	TCTATGCCTC	GGGCATCCAA	GCAGCAAGCG	960
CGTTACGCCG	TGGGTCGATG	TTTGATGTTA	TGGAGCAGCA	ACGATGTTAC	GCAGCAGCAA	1020
CGATGTTACG	CAGCAGGGCA	GTCGCCCTAA	AACAAAGTTA	GGTGGCTCAA	GTATGGGCAT	1080
				ATGCGGGCTG		1140
TTTCGGTCGT	GAGTTCGGAG	ACGTAGCCAC	CTACTCCCAA	CATCAGCCGG	ACTCCGATTA	1200
CCTCGGGAAC	TTGCTCCGTA	GTAAGACATT	CATCGCGCTT	GCTGCCTTCG	ACCAAGAAGC	1260
GGTTGTTGGC	GCTCTCGCGG	CTTACGTTCT	GCCCAGGTTT	GAGCAGCCGC	GTAGTGAGAT	1320
CTATATCTAT	GATCTCGCAG	TCTCCGGCGA	GCACCGGAGG	CAGGGCATTG	CCACCGCGCT	1380
	CTCAAGCATG			TATGTGATCT		1440
				TTGGGCATAC		1500
GATGCACTTT				TTCGTTCAAG		1560
CTTCCCTGAT				GATCAACGGC	ATAAATATTC	1620
	AGGTTTGGCA			TGGACGCCCT	TTGCACGCGC	1680
	TATGCAGCAG		- "		AGCGGGCGGA	1740
	ATGTCTAAAT				CCGTAAACGA	1800
TAAAAATAAA	AAAGAGGAAT				TTCAAAATAG	1860
CTACCCTGGC	AGGCGCCATT	TATGCAGCGC		AGCTGCAAAC	TCGATGCAGC	1920
TGGATGTAGG				ACACCCTCAA		1980
GCCTCTCGCC				CGCCGACTGT		2040
ATCCGGATAT		GGTGATCGTA	ACTTTGACCG	GGGGCTTGGT	ATCCAATCGT	2100
CTCGATATTC	TGGCTGCAG					2119

		i.		 
	i,			
*			~	

CTGCAGCCAG	GGCTGAAAAG	GAGGGATTCA	GTGAGGTCAT	GAAGGGAGGG	GACGGCGCCT	60
GGCTCCAATT	GCTCGATGGC	GCCGCGATTG	AGTGTCTTGG	GCGCGGTCTT	GGAGAGTTCG	120
GCTAGGGAGA	TAAATTTGCT	GGCCATGGTG	GCGGCCCCTG	ATGGGTTGGA	TGATTTTCTG	180
CATTCTGCAT	CATGAAATTC	ATGAAATCAT	CACTTTTCGG	GGGGTGGGTG	CACGGGATTG	240
AAGGTTGCTA	GGAGAGTGCA	TTGCTCGTAA	GCCCAGGAAG	CACGCGGGTT	TCAGGATGGT	300
GCATGGAAAT	GGCATGAGCT	TTGCTGGATA	TGATTAGAGA	CATTAACTAT	TTTGGCGGAA	360
TGGAAGCACG	ATTCCTCGCC	CGGTAGAGCG	GTAACCGCGA	CATTCAGGAC	CGTAAAAAGG	420
AAAGAGCATG	CAACTGACCA	ACAAGAAAAT	CGTCGTCACC	GGAGTGTCCT	CCGGTATCGG	480
TGCCGAAACT	GCCCGCGTTC	TGCGCTCTCA	CGGCGCCACA	GTGATTGGCG	TAGATCGCAA	540
CATGCCGAGC	CTGACTCTGG	ATGCTTTCGT	TCAGGCTGAC	CTGAGCCATC	CTGAAGGCAT	600
CGATCAACGG	CATAAATATT	CCAGTGGACG	GAGGTTTGGC	ATCGACCTAC	GTGTAAGTTC	660
GTGGACGCCC	TTTGCACGCG	CACTATATCT	CTATGCAGCA	GCTGAAAGCA	GCTTTGGTTT	720
TGATCGGAGG	TAGCGGGCGG	AAAGGTGCAG	AATGTCTAAA	TAATAAAGGA	TTCTTGTGAA	780
GCTTTAGTTG	TCCGTAAACG	AAAATAAAA	TAAAGAGGAA	TGATATGAAA	GCAAGTAGAT	840
CAGTCTGCAC	TTTCAAAATA,	GCTACCCTGG	CAGGCGCCAT	TTATGCAGCG	CTGCCAATGT	900
CAGCTGCAAA	CTCGATGCAG	CTGGATGTAG	GTAGCTCGGA	TTGGACGGTG	CGTTGGGGAC	960
AACACCCTCA	AGTATAGCCT	TGCCTCTCGC	CTGAATGAGC	AAGACTCAAG	TCTGACAAAT	1020
GCGCCGACTG	TCAATGGTTA	TATCCGGATA	TTCAAAGTCA	GGGTGATCGT	AACTTTGACC	1080
GGGGGCTTGG	TATCCAATCG	TCTCGATATT	CTGGCTGCAG			1120

			-
		• 	

_						
GAATTCCGCG	TATCGCCCGG	TTCTATCAGC	GGGCCGCTTT	CGAAAGTCAT	GGTGTTAGCC	60
GGTAGGGTCT	TTTTCTTGGC	CATGCTTGTT	GCCTGAACCT	TCGTTGACAT	AGGGCAGAGG	120
TGCGTTTGCC	GCTTCGCTTC	GCGATGAACC	GCATCGAGAT	GCTGAGGTCA	GGATTTTTCC	180
TTAACTCGCG	TAAGCATTCT	GTCATTTTTT	TGGTGGCTTT	GAACAGCCTG	ATGAAAGGTG	240
GTCTCGCCCT	TTGAGGCCGA	TTCTTGGGCG	CTTGGCGGCG	TCGAAGCGAT	GCTCCACTAC	300
CGATTAAGAT	ATTAAAATTAA	AGGAAACCGC	ATGGTTTCTT	ATGTGAATTT	GTCTGGCATA	360
CTCCAGCTCA	AGGGCAATTT	TTGGGCTATT	GGCTGAGCAG	TTGCCTCTAT	ATGGTTATTC	420
AGAATAACAA	TTGACTCCTC	AGGAGGTCAG	CGATGAGCAT	TCTTGGTTTG	AATGGTGCCC	480
CGGTCGGAGC	TGAGCAGCTG	GGCTCGGCTC	TTGATCGCAT	GAAGAAGGCG	CACCTGGAGC	540
AGGGGCCTGC	AAACTTGGAG	CTGCGTCTGA	GTAGGCTGGA	TCGTGCGATT	GCAATGCTTC	600
TGGAAAATCG	TGAAGCAATT	GCCGACGCGG	TTTCTGCTGA	CTTTGGCAAT	CGCAGCCGTG	660
AGCAAACACT	GCTTTGCGAC	ATTGCTGGCT	CGGTGGCAAG	CCTGAAGGAT	AGCCGCGAGC	720
ACGTGGCCAA	ATGGATGGAG	CCCGAACATC	ACAAGGCGAT	GTTTCCAGGG	GCGGAGGCAC	780
GCGTTGAGTT	TCAGCCGCTG	GGTGTCGTTG	GGGTCATTAG	TCCCTGGAAC	TTCCCTATCG	840
TACTGGCCTT	TGGGCCGCTG	GCCGGCATAT	TCGCAGCAGG	TAATCGCGCC	ATGCTCAAGC	900
CGTCCGAGCT	TACCCCGCGG	ACTTCTGCCC	TGCTTGCGGA	GCTAATTGCT	CGTTACTTCG	960
			ACGCTGAAGT	CGGTGCGCTG	TTCAGTGCTC	1020
		TTCACCGGCG			ATCATGCGTG	1080
	TAACCTAGTG		TGGAATTGGG			1140
			CACAACGGGT			1200
			ATGTGCTGCT			1260
			GTAAGGTTGG			1320
			GGCGCAGGGG			1380
			AAGATGGATT		TCTCCGGCCG	1440
			GGGCACAACA			1500
			GCCCGGTTCT			1560
			CAGCGCGGCT			1620
			TCACTGAAGC			1680
			CATCTCACCT			1740
			ATACGCTTGA			1800
			CACGTACTCG			1860
			GGCTCGCGCC			1920
			TCGTCGTGAC			1980
			CTGGATTCAT			2040
			CTACCCGTGA			2100
GCGAATGGGC	TGACCGCTTC	CTCGTGCTTT	ACGGTATCGC	CGCTCCCGAT	TCGCAGCGCA	2160

×.			-
·			
		<b>-</b>	

TCGCCTTCTA	TCGCCTTCTT	GACGAGTTCT	TCTGAGCGGG	ACTCTGGGGT	TCGAAATGAC	2220
CGACCAAGCG	ACGCCCGCCA	TGCCAAGCCT.	GTTCTCGTGC	AAAGTCCTGT	GGGTGAGTCG	2280
AACTTGGCGA	TGCGCGCACC	CTACGGAGAA	GCGATCCACG	GACTGCTCTC	TGTCCTCCTT	2340
TCAACGGAGT	GTTAGAACCG	TTGGTAGTGG	TTTTGGACGG	GCCCAGGAGC	ATGCGCTTCT	2400
GGGCCCGTTT	CTTGAGTATT	CATTGGATAG	TCACGCGTGG	TAGCTTCGAG	CCTGCACAGC	2460
TGATGAGCAC	CCTGGAAGGC	GCGCTGTACG	CGGACGACTG	GGTTCATCTT	CGCCATTCAT	2520
GACGGAACTC	CGTTCCCCAG	TACCGCGATG	ACTATTTTGC	CTCTTCCGAT	GTCCGATTCC	2580
ACGCCGCCTG	ACGCTAAGCG	GGGGCGGGG	CGCCCGCATC	CCAGCCCAGA	CAGCAACAAA	2640
TGAGTAGGCT	CTTGGATGCC	GCGGCGGCTG	AGATTGGTAA	CGGCAATTTC	GTCAATGTGA	2700
CGATGGATTC	GATTGCCCGT	GCTGCCGGCG	TCTCAAAAAA	AACGCTGTAC	GTCTTGGTGG	2760
CGAGCAAGGA	AGAACTCATT	TCCCGGTTAG	TGGCTCGAGA	CATGTCCAAC	CTTGAGGAAT	2820
TC						2822

				-
4:		d)		
·				
			_	
				:

GAATTCCGCG	TATCGCCCGG	TTCTATCAGC	GGGCCGCTTT	CGAAAGTCAT	GGTGTTAGCC	60
GGTAGGGTCT	TTTTCTTGGC	CATGCTTGTT	GCCTGAACCT	TCGTTGACAT	AGGGCAGAGG	120
TGCGTTTGCC	GCTTCGCTTC	GCGATGAACC	GCATCGAGAT	GCTGAGGTCA	GGATTTTTCC	180
TTAACTCGCG	TAAGCATTCT	GTCATTTTTT	TGGTGGCTTT	GAACAGCCTG	ATGAAAGGTG	240
GTCTCGCCCT	TTGAGGCCGA	TTCTTGGGCG	CTTGGCGGCG	TCGAAGCGAT	GCTCCACTAC	300
CGATTAAGAT	ATTAAAATA	AGGAAACCGC	ATGGTTTCTT	ATGTGAATTT	GTCTGGCATA	360
CTCCAGCTCA	AGGGCAATTT	TTGGGCTATT	GGCTGAGCAG	TTGCCTCTAT	ATGGTTATTC	420
AGAATAACAA	TTGACTCCTC	AGGAGGTCAG	CGATGAGCAT	TCTTGGTTTG	AATGGTGCCC	480
CGGTCGGAGC	TGAGCAGCTG	GGCTCGGCTC	TTGATCGCAT	GAAGAAGGCG	CACCTGGAGC	540
AGGGGCCTGC	AAACTTGGAG	CTGCGTCTGA	GTAGGCTGGA	TCGTGCGATT	GCAATGCTTC	600
TGGAAAATCG	TGAAGCAATT	GCCGACGCGG	TTTCTGCTGA	CTTTGGCAAT	CGCAGCCGTG	660
				CCTGAAGGAT		720
ACGTGGCCAA	ATGGATGGAG	CCCGAACATC	ACAAGGCGAT	GTTTCCAGGG	GCGGAGGCAC	780
GCGTTGAGTT	TCAGCCGCTG	GGTGTCGTTG	GGGTCATTAG	TCCCTGGAAC	TTCCCTATCG	840
TACTGGCCTT	TGGGCCGCTG	GCCGGCATAT	TCGCAGCAGG	TAATCGCGCC	ATGCTCAAGC	900
CGTCCGAGCT	TACCCCGCGG	ACTTCTGCCC	TGCTTGCGGA	GCTAATTGCT	CGTTACTTCG	960
ATGAAACTGA	GCTGACTACA	GTGCTGGGCG	ACGCTGAAGT	CGGTGCGCTG	TTCAGTGCTC	1020
AGCCTTTCGA	TCATCTGATC	TTCACCGGCG	GCACTGCCGT	GGCCAAGCAC	ATCATGCGTG	1080
CCGCGGCGGA	TAACCTAGTG	CCCGTTACCC	TGGAATTGGG	TGGCAAATCG	CCGGTGATCG	1140
TTTCCCGCAG	TGCAGATATG	GCGGACGTTG	CACAACGGGT	GTTGACGGTG	AAAACCTTCA	1200
				GGAGAGGCGG		1260
GGGCGCATGC	ATAAAAACTG	TTGTAATTCA	TTAAGCATTC	TGCCGACATG	GAAGCCATCA	1320
				CCTTGTCGCC		1380
TATTTGCCCA	TGGACGCACA	CCGTGGAAAC	GGATGAAGGC	ACGAACCCAG	TTGACATAAG	1440
CCTGTTCGGT	TCGTAAACTG	TAATGCAAGT	AGCGTATGCG	CTCACGCAAC	TGGTCCAGAA	1500
				TTTCATGGCT		1560
				AGCGCGTTAC		1620
				GCAACGATGT		1680
				GCATCATTCG		1740
				ATCTTTTCGG		1800
				ATTACCTCGG		1860
				AAGCGGTTGT		1920
<b></b>	-			AGATCTATAT		1980
				CGCTCATCAA		2040
				AAGCAGATTA		2100
CCCGCAGTGG	CTCTCTATAC	AAAGTTGGGC	ATACGGGAAG	AAGTGATGCA	CTTTGATATC	2160

		_
	#F)	•
		-
		• • :
-		

GACCCAAGTA	CCGCCACCTA	ACAATTCGTT	CAAGCCGAGA	TCGGCTTCCC	TGCAAAGTCC	2220
TGTGGGTGAG	TCGAACTTGG	CGATGCGCGC	ACCCTACGGA	GAAGCGATCC	ACGGACTGCT	2280
CTCTGTCCTC	CTTTCAACGG	AGTGTTAGAA	CCGTTGGTAG	TGGTTTTGGA	CGGGCCCAGG	2340
AGCATGCGCT	TCTGGGCCCG	TTTCTTGAGT	ATTCATTGGA	TAGTCACGCG	TGGTAGCTTC	2400
GAGCCTGCAC	AGCTGATGAG	CACCCTGGAA	GGCGCGCTGT	ACGCGGACGA	CTGGGTTCAT	2460
CTTCGCCATT	CATGACGGAA	CTCCGTTCCC	CAGTACCGCG	ATGACTATTT	TGCCTCTTCC	2520
GATGTCCGAT	TCCACGCCGC	CTGACGCTAA	GCGGGGGCGG	GGGCGCCGC	ATCCCAGCCC	2580
AGACAGCAAC	AAATGAGTAG	GCTCTTGGAT	GCCGCGGCGG	CTGAGATTGG	TAACGGCAAT	2640
TTCGTCAATG	TGACGATGGA	TTCGATTGCC	CGTGCTGCCG	GCGTCTCAAA	AAAAACGCTG	2700
TACGTCTTGG	TGGCGAGCAA	GGAAGAACTC	ATTTCCCGGT	TAGTGGCTCG	AGACATGTCC	2760
AACCTTGAGG	AATTC					2775

		-
.*., 	<b>-</b>	

GAATTCCGCG	TATCGCCCGG	TTCTATCAGC	GGGCCGCTTT	CGAAAGTCAT	GGTGTTAGCC	60
GGTAGGGTCT	TTTTCTTGGC	CATGCTTGTT	GCCTGAACCT	TCGTTGACAT	AGGGCAGAGG	120
TGCGTTTGCC	GCTTCGCTTC	GCGATGAACC	GCATCGAGAT	GCTGAGGTCA	GGATTTTTCC	180
TTAACTCGCG	TAAGCATTCT	GTCATTTTT	TGGTGGCTTT	GAACAGCCTG	ATGAAAGGTG	240
GTCTCGCCCT	TTGAGGCCGA	TTCTTGGGCG	CTTGGCGGCG	TCGAAGCGAT	GCTCCACTAC	300
CGATTAAGAT	ATTAAAATA	AGGAAACCGC	ATGGTTTCTT	ATGTGAATTT	GTCTGGCATA	360
CTCCAGCTCA	AGGGCAATTT	TTGGGCTATT	GGCTGAGCAG	TTGCCTCTAT	ATGGTTATTC	420
AGAATAACAA	TTGACTCCTC	AGGAGGTCAG	CGATGAGCAT	TCTTGGTTTG	AATGGTGCCC	480
CGGTCGGAGC	TGAGCAGCTG	GGCTCGGCTC	TTGATCGCAT	GAAGAAGGCG	CACCTGGAGC	540
AGGGGCCTGC	AAACTTGGAG	CTGCGTCTGA	GTAGGCTGGA	TCGTGCGATT	GCAATGCTTC	600
TGGAAAATCG	TGAAGCAATT	GCCGACGCGG	TTTCTGCTGA	CTTTGGCAAT	CGCAGCCGTG	660
AGCAAACACT	GCTTTGCGAC	ATTGCTGGCT	CGGTGGCAAG	CCTGAAGGAT	AGCCGCGAGC	720
ACGTGGCCAA	ATGGATGGAG	CCCGAACATC	ACAAGGCGAT	GTTTCCAGGG	GCGGAGGCAC	780
GCGTTGAGTT	TCAGCCGCTG	GGTGTCGTTG	GGGTCATTAG	TCCCTGGAAC	TTCCCTATCG	840
TACTGGCCTT	TGGGCCGCTG	GCCGGCATAT	TCGCAGCAGG	TAATCGCGCC	ATGCTCAAGC	900
CGTCCGAGCT	TACCCCGCGG	ACTTCTGCCC	TGCTTGCGGA	GCTAATTGCT	CGTTACTTCG	960
ATGAAACTGA	GCTGACTACA	GTGCTGGGCG	ACGCTGAAGT	CGGTGCGCTG	TTCAGTGCTC	1020
AGCCTTTCGA	TCATCTGATC	TTCACCGGCG	GCACTGCCGT	GGCCAAGCAC	ATCATGCGTG	1080
CCGCGGCGGA	TAACCTAGTG	CCCGTTACCC	TGGAATTGGG	TGGCAAATCG	CCGGTGATCG	1140
TTTCCCGCAG	TGCAGATATG	GCGGACGTTG	CACAACGGGT	GTTGACGGTG	AAAACCTTCA	1200
ATGCCGGGCA	AATCTGTCTG	GCACCGTGGG	TGAGTCGAAC	TTGGCGATGC	GCGCACCCTA	1260
CGGAGAAGCG	ATCCACGGAC	TGCTCTCTGT	CCTCCTTTCA	ACGGAGTGTT	AGAACCGTTG	1320
GTAGTGGTTT	TGGACGGGCC	CAGGAGCATG	CGCTTCTGGG	CCCGTTTCTT	GAGTATTCAT	1380
TGGATAGTCA	CGCGTGGTAG	CTTCGAGCCT	GCACAGCTGA	TGAGCACCCT	GGAAGGCGCG	1440
CTGTACGCGG	ACGACTGGGT	TCATCTTCGC	CATTCATGAC	GGAACTCCGT	TCCCCAGTAC	1500
CGCGATGACT	ATTTTGCCTC	TTCCGATGTC	CGATTCCACG	CCGCCTGACG	CTAAGCGGGG	1560
GCGGGGGCGC	CCGCATCCCA	GCCCAGACAG	CAACAAATGA	GTAGGCTCTT	GGATGCCGCG	1620
GCGGCTGAGA	TTGGTAACGG	CAATTTCGTC	AATGTGACGA	TGGATTCGAT	TGCCCGTGCT	1680
GCCGGCGTCT	СААААААААС	GCTGTACGTC	TTGGTGGCGA	GCAAGGAAGA	ACTCATTTCC	1740
CGGTTAGTGG	CTCGAGACAT	GTCCAACCTT	GAGGAATTC			1779

				4	
					-
· .					
			_		

60

CTGCAGCCGA	GCATCGATTG			TGGCTGACCA		60
CCCGCGGCAC	TATCCAATCT		TCGGGCGCCG	CGGGCATCAT	GCCCGCGGCG	120
CTCGCCTCAT	TTCAATCTCT			GTTCTCCGGT	CTTGGTGGAT	180
CAAGGCCAGT	CGCGGAGAGT	CTCGAAGAGG	AGAGTACAGT	GAACGCCGAG	TCCACATTGC	240
AACCGCAGGC	ATCATCATGC	TCTGCTCAGC	CACGCTACCG	CAGTGTGTCG	ATTGGTCATC	300
CTCCGGTTGA	GGTTACGCAA	GACGCTGGAG	GTATTGTCCG	GATGCGTTCT	CTCGAGGCGC	360
TTCTTCCCTT	CCCGGGTCGA		GTCTCGAGCA	TTGGGCTAAG		420
AACAAACCTG	CGTTGCTGCC	AGGGCGGCAA	ATGGGGAATG	GCGTCGTATC	AGCTACGCGG	480
AAATGTTCCA	CAACGTCCGC	GCCATCGCAC	AGAGCTTGCT	TCCTTACGGA	CTATCGGCAG	540
AGCGTCCGCT	GCTTATCGTC	TCTGGAAATG	ACCTGGAACA	TCTTCAGCTG	GCATTTGGGG	600
CTATGTATGC	GGGCATTCCC	TATTGCCCGG	TGTCTCCTGC	TTATTCACTG	CTGTCGCAAG	660
ATTTGGCGAA	GCTGCGTCAC	ATCGTAGGTC	TTCTGCAACC	GGGACTGGTC	TTTGCTGCCG	720
ATGCAGCACC	TTTCCAGGGG	ACAGCAAGCG	AACCGGAATT	GCCAGCTGGG	GCGCCCTCTG	780
GTAAGGTTGG	GAAGCCCTGC	AAAGTAAACT	GGATGGCTTT	CTTGCCGCCA	AGGATCTGAT	840
GGCGCAGGGG	ATCAAGATCT	GATCAAGAGA	CAGGATGAGG	ATCGTTTCGC	ATGATTGAAC	900
AAGATGGATT	GCACGCAGGT	TCTCCGGCCG	CTTGGGTGGA	GAGGCTATTC	GGCTATGACT	960
GGGCACAACA	GACAATCGGC	TGCTCTGATG	CCGCCGTGTT	CCGGCTGTCA	GCGCAGGGGC	1020
GCCCGGTTCT	TTTTGTCAAG	ACCGACCTGT	CCGGTGCCCT	GAATGAACTG	CAGGACGAGG	1080
CAGCGCGGCT	ATCGTGGCTG	GCCACGACGG	GCGTTCCTTG	CGCAGCTGTG	CTCGACGTTG	1140
TCACTGAAGC	GGGAAGGGAC	TGGCTGCTAT	TGGGCGAAGT	GCCGGGGCAG	GATCTCCTGT	1200
CATCTCACCT	TGCTCCTGCC	GAGAAAGTAT	CCATCATGGC	TGATGCAATG	CGGCGGCTGC	1260
ATACGCTTGA	TCCGGCTACC	TGCCCATTCG	ACCACCAAGC	GAAACATCGC	ATCGAGCGAG	1320
CACGTACTCG	GATGGAAGCC	GGTCTTGTCG	ATCAGGATGA	TCTGGACGAA	GAGCATCAGG	1380
GGCTCGCGCC	AGCCGAACTG	TTCGCCAGGC	TCAAGGCGCG	CATGCCCGAC	GGCGAGGATC	1440
TCGTCGTGAC	CCATGGCGAT	GCCTGCTTGC	CGAATATCAT	GGTGGAAAAT	GGCCGCTTTT	1500
CTGGATTCAT	CGACTGTGGC	CGGCTGGGTG	TGGCGGACCG	CTATCAGGAC	ATAGCGTTGG	1560
CTACCCGTGA	TATTGCTGAA	GAGCTTGGCG	GCGAATGGGC	TGACCGCTTC	CTCGTGCTTT	1620
ACGGTATCGC	CGCTCCCGAT	TCGCAGCGCA	TCGCCTTCTA	TCGCCTTCTT	GACGAGTTCT	1680
TCTGAGCGGG	ACTCTGGGGT	TCGAAATGAC	CGACCAAGCG	ACGCCCCTGT	TTTGCAATGG	1740
CGGTCGGCGA	AAGTTGATGC	GCTGTATCGT	GGTGAAGATC	AATCCATGCT	GCGTGACGAG	1800
GCCACACTGT	GAGTTGGTCA	GGGGGGCTT	ACTCGGCGTT	TTCCGACACT	GCGTTGGTTG	1860
CGGCAGTGCG	CACCCCCTGG	ATTGATTGCG	GGGGTGCCCT	GTCGCTGGTG	TCGCCTATCG	1920
ACTTAGGGGT	AAAGGTCGCT	CGCGAAGTTC	TGATGCGTGC	GTCGCTTGAA	CCACAAATGG	1980
TCGATAGCGT	ACTCGCAGGC	TCTATGGCTC	AAGCAAGCTT	TGATGCTTAC	CTGCTCCCGC	2040
GGCACATTGG	CTTGTACAGC	GGTGTTCCCA	AGTCGGTTCC	GGCCTTGGGG	GTGCAGCGCA	2100
TTTGCGGCAC	AGGCTTCGAA	CTGCTTCGGC	AGGCCGGCGA	GCAGATTTCC	CAAGGCGCTG	2160
ATCACGTGCT	GTGTGTCGCG	GGCTGCAG				2188

			. · · · · ·
			-
	·		
ig:		ţ	
,			
		<b>-</b>	

CTGCAGCCGA	GCATCGATTG	AGCACTTTAC	CCAGCTGCGC	TGGCTGACCA	TTCAGAATGG	60
CCCGCGGCAC	TATCCAATCT	AAATCGATCT	TCGGGCGCCG	CGGGCATCAT	GCCCGCGGCG	120
CTCGCCTCAT	TTCAATCTCT	AACTTGATAA	AAACAGAGCT	GTTCTCCGGT	CTTGGTGGAT	180
CAAGGCCAGT	CGCGGAGAGT	CTCGAAGAGG	AGAGTACAGT	GAACGCCGAG	TCCACATTGC	240
AACCGCAGGC		TCTGCTCAGC	CACGCTACCG	CAGTGTGTCG	ATTGGTCATC	300
	GGTTACGCAA		GTATTGTCCG		CTCGAGGCGC	360
TTCTTCCCTT	CCCGGGTCGA	ATTCTTGAGC	GTCTCGAGCA	TTGGGCTAAG	ACCCGTCCAG	420
AACAAACCTG	CGTTGCTGCC	AGGGCGCAA	ATGGGGAATG	GCGTCGTATC	AGCTACGCGG	480
	CAACGTCCGC		AGAGCTTGCT		CTATCGGCAG	540
AGCGTCCGCT	GCTTATCGTC	TCTGGAAATG	ACCTGGAACA	TCTTCAGCTG	GCATTTGGGG	600
CTATGTATGC		TATTGCCCGG	TGTCTCCTGC	TTATTCACTG	CTGTCGCAAG	660
ATTTGGCGAA	GCTGCGTCAC	ATCGTAGGTC	TTCTGCAACC	GGGACTGGTC	TTTGCTGCCG	720
ATGCAGCACC	TTTCCAGGGG	GAGAGGCGGT	TTGCGTATTG	GGCGCATGCA	TAAAAACTGT	780
TGTAATTCAT	TAAGCATTCT	GCCGACATGG	AAGCCATCAC	AAACGGCATG	ATGAACCTGA	840
ATCGCCAGCG	GCATCAGCAC	CTTGTCGCCT	TGCGTATAAT	ATTTGCCCAT	GGACGCACAC	900
CGTGGAAACG	GATGAAGGCA	CGAACCCAGT	TGACATAAGC	CTGTTCGGTT	CGTAAACTGT	960
AATGCAAGTA	GCGTATGCGC	TCACGCAACT	GGTCCAGAAC	CTTGACCGAA	CGCAGCGGTG	1020
GTAACGGCGC	AGTGGCGGTT	TTCATGGCTT	GTTATGACTG	TTTTTTTGTA	CAGTCTATGC	1080
CTCGGGCATC	CAAGCAGCAA	GCGCGTTACG	CCGTGGGTCG	ATGTTTGATG	TTATGGAGCA	1140
GCAACGATGT	TACGCAGCAG	CAACGATGTT	ACGCAGCAGG	GCAGTCGCCC	TAAAACAAAG	1200
TTAGGTGGCT	CAAGTATGGG	CATCATTCGC	ACATGTAGGC	TCGGCCCTGA	CCAAGTCAAA	1260
TCCATGCGGG	CTGCTCTTGA	TCTTTTCGGT	CGTGAGTTCG	GAGACGTAGC	CACCTACTCC	1320
CAACATCAGC	CGGACTCCGA	TTACCTCGGG	AACTTGCTCC	GTAGTAAGAC	ATTCATCGCG	1380
CTTGCTGCCT	TCGACCAAGA		GGCGCTCTCG		TCTGCCCAGG	1440
	CGCGTAGTGA		TATGATCTCG	CAGTCTCCGG	CGAGCACCGG	1500
	TTGCCACCGC				0000011001	1560
GCTTATGTGA	TCTACGTGCA	AGCAGATTAC	GGTGACGATC		TCTCTATACA	1620
	TACGGGAAGA		TTTGATATCG	· · · · · · · · ·	CGCCACCTAA	1680
	AAGCCGAGAT		TGTTTTGCAA	TGGCGGTCGG	CGAAAGTTGA	1740
	CGTGGTGAAG		GCTGCGTGAC		TGTGAGTTGG	1800
	CTTACTCGGC		ACTGCGTTGG	TTGCGGCAGT	GCGCACCCC	1860
TGGATTGATT	GCGGGGGTGC		GTGTCGCCTA	TCGACTTAGG	GGTAAAGGTC	1920
GCTCGCGAAG	TTCTGATGCG	TGCGTCGCTT	GAACCACAAA	TGGTCGATAG	CGTACTCGCA	1980
GGCTCTATGG	CTCAAGCAAG		TACCTGCTCC	CGCGGCACAT	TGGCTTGTAC	2040
	CCAAGTCGGT	TCCGGCCTTG	GGGGTGCAGC		CACAGGCTTC	2100
GAACTGCTTC	GGCAGGCCGG	CGAGCAGATT	TCCCAAGGCG	CTGATCACGT	GCTGTGTGTC	2160
GCGGGCTGCA	G					2171

		•
		•
		•
*		· ·
•	<b>-</b>	

CTGCAGCCGA	GCATCGATTG	AGCACTTTAC	CCAGCTGCGC	TGGCTGACCA	TTCAGAATGG	60
CCCGCGGCAC	TATCCAATCT	AAATCGATCT	TCGGGCGCCG	CGGGCATCAT	GCCCGCGGCG	120
CTCGCCTCAT	TTCAATCTCT	<b>AACTTGATAA</b>	AAACAGAGCT	GTTCTCCGGT	CTTGGTGGAT	180
CAAGGCCAGT	CGCGGAGAGT	CTCGAAGAGG	AGAGTACAGT	GAACGCCGAG	TCCACATTGC	240
AACCGCAGGC	ATCATCATGC	TCTGCTCAGC	CACGCTACCG	CAGTGTGTCG	ATTGGTCATC	300
CTCCGGTTGA	GGTTACGCAA	GACGCTGGAG	GTATTGTCCG	GATGCGTTCT	CTCGAGGCGC	360
TTCTTCCCTT	CCCGGGTCGA	ATTCTTGAGC	GTCTCGAGCA	TTGGGCTAAG	ACCCGTCCAG	420
AACAAACCTG	CGTTGCTGCC	AGGGCGGCAA	ATGGGGAATG	GCGTCGTATC	AGCTACGCGG	480
AAATGTTCCA	CAACGTCCGC	GCCATCGCAC	AGAGCTTGCT	TCCTTACGGA	CTATCGGCAG	540
AGCGTCCGCT	GCTTATCGTC	TCTGGAAATG	ACCTGGAACA	TCTTCAGCTG	GCATTTGGGG	600
CTATGTATGC	GGGCATTCCC	TATTGCCCGG	TGTCTCCTGC	TTATTCACTG	CTGTCGCAAG	660
ATTTGGCGAA	GCTGCGTCAC	ATCGTAGGTC	TTCTGCAACC	GGGACTGGTC	TTTGCTGCCG	720
ATGCAGCACC	TTTCCAGCGC	GCTGTTTTGC	AATGGCGGTC	GGCGAAAGTT	GATGCGCTGT	780
ATCGTGGTGA	AGATCAATCC	ATGCTGCGTG	ACGAGGCCAC	ACTGTGAGTT	GGTCAGGGGG	840
GGCTTACTCG	GCGTTTTCCG	ACACTGCGTT	GGTTGCGGCA	GTGCGCACCC	CCTGGATTGA	900
TTGCGGGGGT	GCCCTGTCGC	TGGTGTCGCC	TATCGACTTA	GGGGTAAAGG	TCGCTCGCGA	960
AGTTCTGATG	CGTGCGTCGC	TTGAACCACA	AATGGTCGAT	AGCGTACTCG	CAGGCTCTAT	1020
GGCTCAAGCA	AGCTTTGATG	CTTACCTGCT	CCCGCGGCAC	ATTGGCTTGT	ACAGCGGTGT	1080
TCCCAAGTCG	GTTCCGGCCT	TGGGGGTGCA	GCGCATTTGC	GGCACAGGCT	TCGAACTGCT	1140
TCGGCAGGCC	GGCGAGCAGA	TTTCCCAAGG	CGCTGATCAC	GTGCTGTGTG	TCGCGGGCTG	1200
CAG						1203

				, ,
				-
				•
• • • • • • • • • • • • • • • • • • •	· .,	د. مدر		
			_	*

GAATTCCCCT	GGCGACGAAA	GGGCGGCAGG	CCGCATGGCC	ACGGCTGGGC	GGTAACTGAT	60
GCTTGCGTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA	TTTCGGCGAG	AGAATCATGC	120
GGGTACGCCT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC	CTTGAATCAG	AAAAATAGTT	180
AATTGACAGA	ACTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG	240
GGTGCACGAT	GAATAGCTAC	GATGGCCGTT	GGTCTACCGT	TGATGTGAAG	GTTGAAGAAG	300
GTATCGCTTG	GGTCACGCTG	AACCGCCCGG	AGAAGCGCAA	CGCAATGAGC	CCAACTCTCA	360
ATCGAGAGAT	GGTCGAGGTT	CTGGAGGTGC	TGGAGCAGGA	CGCAGATGCT	CGCGTGCTTG	420
TTCTGACTGG	TGCAGGCGAA	TCCTGGACCG	CGGGCATGGA	CCTGAAGGAG	TATTTCCGCG	480
AGACCGATGC	TGGCCCCGAA	ATTCTGCAAG	AGAAGATTCG	TCGGGGACAG	CAAGCGAACC	540
GGAATTGCCA	GCTGGGGCGC	CCTCTGGTAA	GGTTGGGAAG	CCCTGCAAAG	TAAACTGGAT	600
GGCTTTCTTG	CCGCCAAGGA	TCTGATGGCG	CAGGGGATCA	AGATCTGATC	AAGAGACAGG	660
ATGAGGATCG	TTTCGCATGA	TTGAACAAGA	TGGATTGCAC	GCAGGTTCTC	CGGCCGCTTG	720
GGTGGAGAGG	CTATTCGGCT	ATGACTGGGC	ACAACAGACA	ATCGGCTGCT	CTGATGCCGC	780
CGTGTTCCGG	CTGTCAGCGC	AGGGGCGCCC	GGTTCTTTTT	GTCAAGACCG	ACCTGTCCGG	840
TGCCCTGAAT	GAACTGCAGG	ACGAGGCAGC	GCGGCTATCG	TGGCTGGCCA	CGACGGGCGT	900
TCCTTGCGCA	GCTGTGCTCG	ACGTTGTCAC	TGAAGCGGGA	AGGGACTGGC	TGCTATTGGG	960
CGAAGTGCCG	GGGCAGGATC	TCCTGTCATC	TCACCTTGCT	CCTGCCGAGA	AAGTATCCAT	1020
CATGGCTGAT	GCAATGCGGC	GGCTGCATAC	GCTTGATCCG	GCTACCTGCC	CATTCGACCA	1080
CCAAGCGAAA	CATCGCATCG	AGCGAGCACG	TACTCGGATG		TTGTCGATCA	1140
GGATGATCTG	GACGAAGAGC	ATCAGGGGCT	CGCGCCAGCC	GAACTGTTCG	CCAGGCTCAA	1200
GGCGCGCATG	CCCGACGGCG	AGGATCTCGT	CGTGACCCAT	GGCGATGCCT	GCTTGCCGAA	1260
TATCATGGTG	GAAAATGGCC	GCTTTTCTGG	ATTCATCGAC	TGTGGCCGGC	TGGGTGTGGC	1320
GGACCGCTAT	CAGGACATAG	CGTTGGCTAC	CCGTGATATT	GCTGAAGAGC	TTGGCGGCGA	1380
	CGCTTCCTCG		TATCGCCGCT	· · · · · · · - · - <del>-</del>	AGCGCATCGC	1440
			AGCGGGACTC			1500
				GAGAAAAGCA		1560
				CGCTGCGCCC		1620
				CCCTGCTTAT		1680
				GCCCGCTGAC		1740
				CCGCAGTGGC		1800
				GCCGTGCCCG		1860
				CCGCAGCGAG		1920
	ACTGGTATGG	GTTTAACGTT	TACCTGGCGG	CGGGCATGTT	GCGGGGAATT	1980
С						1981

			,	
			4	
			•	
			•	
			4	
	4			
4				
			,	
* . *		••• ••		

	GGCGACGAAA					60
GCTTGCGTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA	TTTCGGCGAG	AGAATCATGC	120
GGGTACGCCT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC	CTTGAATCAG	AAAAATAGTT	180
AATTGACAGA	ACTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG	240
GGTGCACGAT	GAATAGCTAC	GATGGCCGTT	GGTCTACCGT	TGATGTGAAG	GTTGAAGAAG	300
GTATCGCTTG	GGTCACGCTG	AACCGCCCGG	AGAAGCGCAA	CGCAATGAGC	CCAACTCTCA	360
ATCGAGAGAT	GGTCGAGGTT	CTGGAGGTGC	TGGAGCAGGA	CGCAGATGCT	CGCGTGCTTG	420
TTCTGACTGG	TGCAGGCGAA	TCCTGGACCG	CGGGCATGGA	CCTGAAGGAG	TATTTCCGCG	480
AGACCGATGC	TGGCCCCGAA	ATTCTGCAAG	AGAAGATTCG	TCGGGGGAGA	GGCGGTTTGC	540
GTATTGGGCG	CATGCATAAA	AACTGTTGTA	ATTCATTAAG	CATTCTGCCG	ACATGGAAGC	600
CATCACAAAC	GGCATGATGA	ACCTGAATCG	CCAGCGGCAT	CAGCACCTTG	TCGCCTTGCG	660
TATAATATT	GCCCATGGAC	GCACACCGTG	GAAACGGATG	AAGGCACGAA	CCCAGTTGAC	720
ATAAGCCTGT	TCGGTTCGTA	AACTGTAATG	CAAGTAGCGT	ATGCGCTCAC	GCAACTGGTC	780
CAGAACCTTG	ACCGAACGCA	GCGGTGGTAA	CGGCGCAGTG	GCGGTTTTCA	TGGCTTGTTA	840
TGACTGTTTT	TTTGTACAGT	CTATGCCTCG	GGCATCCAAG	CAGCAAGCGC	GTTACGCCGT	900
GGGTCGATGT	TTGATGTTAT	GGAGCAGCAA	CGATGTTACG	CAGCAGCAAC	GATGTTACGC	960
AGCAGGGCAG	TCGCCCTAAA	ACAAAGTTAG	GTGGCTCAAG	TATGGGCATC	ATTCGCACAT	1020
GTAGGCTCGG	CCCTGACCAA	GTCAAATCCA	TGCGGGCTGC	TCTTGATCTT	TTCGGTCGTG	1080
AGTTCGGAGA	CGTAGCCACC	TACTCCCAAC	ATCAGCCGGA	CTCCGATTAC	CTCGGGAACT	1140
TGCTCCGTAG	TAAGACATTC	ATCGCGCTTG	CTGCCTTCGA	CCAAGAAGCG	GTTGTTGGCG	1200
CTCTCGCGGC	TTACGTTCTG	CCCAGGTTTG	AGCAGCCGCG	TAGTGAGATC	TATATCTATG	1260
ATCTCGCAGT	CTCCGGCGAG	CACCGGAGGC	AGGGCATTGC	CACCGCGCTC	ATCAATCTCC	1320
	GGCCAACGCG				GATTACGGTG	1380
ACGATCCCGC	AGTGGCTCTC	TATACAAAGT	TGGGCATACG	GGAAGAAGTG	ATGCACTTTG	1440
	AAGTACCGCC				TTCCCCGAGC	1500
	GCAGTTCCTT					1560
	GCGCCGGGGC				ACAATAATGA	1620
GGAGTGCCCA	ATGTTTCACG				CAGCATCTGA	1680
TGAGCGCACC	•		GACCGGAGAA		GCGTCGCTGC	1740
TGCCAGTTTG	GAAGATGCGG				TTCCTGAATG	1800
GGCGGCGCTT				CGAGCGGCGG		1860
GGACCGTTCT				GGCGCAGCGG	GAAACTGGTA	1920
TGGGTTTAAC	GTTTACCTGG	CGGCGGCAT	GTTGCGGGGA	ATTC		1964

			_
			<b>4</b> 1
			•
			•
÷			
			ž.
	•		
			·
		-	<del>,"</del>

GAATTCCCCT	GGCGACGAAA	GGGCGGCAGG	CCGCATGGCC	ACGGCTGGGC	GGTAACTGAT	60
GCTTGCGTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA	TTTCGGCGAG	AGAATCATGC	120
GGGTACGCCT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC	CTTGAATCAG	AAAAATAGTT	180
AATTGACAGA	ACTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG	240
GGTGCACGAT	GAATAGCTAC	GATGGCCGTT	GGTCTACCGT	TGATGTGAAG	GTTGAAGAAG	300
GTATCGCTTG	GGTCACGCTG	AACCGCCCGG	AGAAGCGCAA	CGCAATGAGC	CCAACTCTCA	360
ATCGAGAGAT	GGTCGAGGTT	CTGGAGGTGC	TGGAGCAGGA	CGCAGATGCT	CGCGTGCTTG	420
TTCTGACTGG	TGCAGGCGAA	TCCTGGACCG	CGGGCATGGA	CCTGAAGGAG	TATTTCCGCG	480
AGACCGATGC	TGGCCCCGAA	ATTCTGCAAG	AGAAGATTCG	TCGCGAGCAG	GGCATGAAGC	540
AGTTCCTTGA	CGAGAAAAGC	ATCAAGCCGG	GCTTGCAGAC	CTACAAGCGC	TGATAAATGC	600
GCCGGGGCCC	TCGCTGCGCC	CCCGGCCTTC	CAATAATGAC	AATAATGAGG	AGTGCCCAAT	660
GTTTCACGTG	CCCCTGCTTA	TTGGTGGTAA	GCCTTGTTCA	GCATCTGATG	AGCGCACCTT	720
CGAGCGTCGT	AGCCCGCTGA	CCGGAGAAGT	GGTATCGCGC	GTCGCTGCTG	CCAGTTTGGA	780
AGATGCGGAC	GCCGCAGTGG	CCGCTGCACA	GGCTGCGTTT	CCTGAATGGG	CGGCGCTTGC	840
TCCGAGCGAA	CGCCGTGCCC	GACTGCTGCG	AGCGGCGGAT	CTTCTAGAGG	ACCGTTCTTC	900
CGAGTTCACC	GCCGCAGCGA	GTGAAACTGG	CGCAGCGGGA	AACTGGTATG	GGTTTAACGT	960
TTACCTGGCG	GCGGGCATGT	TGCGGGGAAT	TC			992

-

	• •			
				-
				*
	· )			•
				,
			**:	
		·		
				4
,				•

			CCCAATGTTT			60
TGGTAAGCCT			CACCTTCGAG			120
AGAAGTGGTA			TTTGGAAGAT			180
TGCACAGGCT			GCTTGCTCCG			240
			TTCTTCCGAG			300
			TAACGTTTAC			360
			GGGCGATGTC			420
			TGGCGTGGTG			480
			TGCGATGCCG			540
GGTGTTGAAA	AGCTCTGAGC	TGAGTCCCTT	TACCCATCGC	CTGATTGGTC	AGGTGTTGCA	600
			TGTCATCAGC		AAGACGCTCC	660
TGCGGTGGTG	GAGCGACTGA	TTGCAAATCC	TGCGGTACGT	CGAGTGAACT	TCACCGGTTC	720
GACCCACGTT	GGACGGATCA	TTGGTGAGCT	GTCTGCGCGT	CATCTGAAGC	CTGCTGTGCT	780
GGAATTAGGT	GGTAAGGCTC	CGTTCTTGGT	CTTGGACGAT	GCCGACCTCG	ATGCGGCGGT	840
CGAAGCGGCG	GCCTTTGGTG	CCTACTTCAA	TCAGGGTCAA	ATCTGCATGT	CCACTGAGCG	900
TCTGATTGTG	ACAGCAGTCG	CAGACGCCTT	TGTTGAAAAG	CTGGCGAGGA	AGGTCGCCAC	960
ACTGCGTGCT	GGCGATCCTA	ATGATCCGCA	ATCGGTCTTG	GGTTCGTTGA	TTGATGCCAA	1020
TGCAGGTCAA	CGCATCCAGG	TTCTGGTCGA	TGATGCGCTC	GGGGACAGCA	AGCGAACCGG	1080
AATTGCCAGC	TGGGGCGCCC	TCTGGTAAGG	TTGGGAAGCC	CTGCAAAGTA	AACTGGATGG	1140
CTTTCTTGCC	GCCAAGGATC	TGATGGCGCA	GGGGATCAAG	ATCTGATCAA	GAGACAGGAT	1200
0.100.1100.1			GATTGCACGC			1260
			AACAGACAAT			1320
			TTCTTTTTGT			1380
			GGCTATCGTG			1440
			AAGCGGGAAG			1500
			ACCTTGCTCC			1560
			TTGATCCGGC			1620
			CTCGGATGGA			1680
			CGCCAGCCGA			1740
CGCGCATGCC	CGACGGCGAG	GATCTCGTCG	TGACCCATGG	CGATGCCTGC	TTGCCGAATA	1800
	AAATGGCCGC		TCATCGACTG			1860
			GTGATATTGC			1920
GGGCTGACCG			TCGCCGCTCC			1980
TCTATCGCCT			CGGGACTCTG			2040
			GCATTTGCCA			2100
ATGACGAGGC	TCAGATGCCA	TTCGGTGGGG	TGAAGTCCAG	CGGCTACGGC	AGCTTCGGCA	2160

	7° <del>-</del>				~
		ý.			ėų.
					·
					÷
					,.
					-
	<i>(</i> ************************************		_	•	

GTCGAGCATC	GATTGAGCAC	TTTACCCAGC	TGCGCTGGCT	GACCATTCAG	AATGGCCCGC	2220
GGCACTATCC	AATCTAAATC	GATCTTCGGG	CGCCGCGGGC	ATCATGCCCG	CGGCGCTCGC	2280
CTCATTTCAA	TCTCTAACTT	GATAAAAACA	GAGCTGTTCT	CCGGTCTTGG	TGGATCAAGG	2340
CCAGTCGCGG	AGAGTCTCGA	AGAGGAGAGT	ACAGTGAACG	CCGAGTCCAC	ATTGCAACCG	2400
CAGGCATCAT	CATGCTCTGC	TCAGCCACGC	TACCGCAGTG	TGTCGATTGG	TCATCCTCCG	2460
GTTGAGGTTA	CGCAAGACGC	TGGAGGTATT	GTCCGGATGC	GTTCTCTCGA	GGCGCTTCTT	2520 <sup>-</sup>
CCCTTCCCGG	GTGGAATTC					2539

	9	
		•
*		
		1.
		*

GAATTCCAAT	AATGACAATA	ATGAGGAGTG	CCCAATGTTT	CACGTGCCCC	TGCTTATTGG	60
TGGTAAGCCT	TGTTCAGCAT	CTGATGAGCG	CACCTTCGAG	CGTCGTAGCC	CGCTGACCGG	120
AGAAGTGGTA	TCGCGCGTCG	CTGCTGCCAG	TTTGGAAGAT	GCGGACGCCG	CAGTGGCCGC	180
TGCACAGGCT	GCGTTTCCTG	AATGGGCGGC	GCTTGCTCCG	AGCGAACGCC	GTGCCCGACT	240
GCTGCGAGCG	GCGGATCTTC	TAGAGGACCG	TTCTTCCGAG	TTCACCGCCG	CAGCGAGTGA	300
AACTGGCGCA	GCGGGAAACT	GGTATGGGTT	TAACGTTTAC	CTGGCGGCGG	GCATGTTGCG	360
GGAAGCCGCG	GCCATGACCA	CACAGATTCA	GGGCGATGTC	ATTCCGTCCA	ATGTGCCCGG	420
TAGCTTTGCC	ATGGCGGTTC	GACAGCCATG	TGGCGTGGTG	CTCGGTATTG	CGCCTTGGAA	480
TGCTCCGGTA	ATCCTTGGCG	TACGGGCTGT	TGCGATGCCG	TTGGCATGCG	GCAATACCGT	540
GGTGTTGAAA	AGCTCTGAGC	TGAGTCCCTT	TACCCATCGC	CTGATTGGTC	AGGTGTTGCA	600
TGATGCTGGT	CTGGGGGATG	GCGTGGTGAA	TGTCATCAGC	AATGCCCCGC	AAGACGCTCC	660
TGCGGTGGTG	GAGCGACTGA	TTGCAAATCC	TGCGGTACGT	CGAGTGAACT	TCACCGGTTC	720
GACCCACGTT	GGACGGATCA	TTGGTGAGCT	GTCTGCGCGT	CATCTGAAGC	CTGCTGTGCT	780
GGAATTAGGT	GGTAAGGCTC	CGTTCTTGGT	CTTGGACGAT	GCCGACCTCG	ATGCGGCGGT	840
CGAAGCGGCG	GCCTTTGGTG	CCTACTTCAA	TCAGGGTCAA	ATCTGCATGT	CCACTGAGCG	900
TCTGATTGTG	ACAGCAGTCG	CAGACGCCTT	TGTTGAAAAG	CTGGCGAGGA	AGGTCGCCAC	960
ACTGCGTGCT	GGCGATCCTA	ATGATCCGCA	ATCGGTCTTG	GGTTCGTTGA	TTGATGCCAA	1020
TGCAGGTCAA	CGCATCCAGG	TGGGGAGAGG	CGGTTTGCGT	ATTGGGCGCA	TGCATAAAAA	1080
CTGTTGTAAT	TCATTAAGCA	TTCTGCCGAC	ATGGAAGCCA	TCACAAACGG	CATGATGAAC	1140
CTGAATCGCC	AGCGGCATCA	GCACCTTGTC	GCCTTGCGTA	TAATATTTGC	CCATGGACGC	1200
ACACCGTGGA	AACGGATGAA	GGCACGAACC	CAGTTGACAT	AAGCCTGTTC	GGTTCGTAAA	1260
CTGTAATGCA	AGTAGCGTAT	GCGCTCACGC	AACTGGTCCA	GAACCTTGAC	CGAACGCAGC	1320
GGTGGTAACG	GCGCAGTGGC	GGTTTTCATG	GCTTGTTATG	ACTGTTTTT	TGTACAGTCT	1380
ATGCCTCGGG	CATCCAAGCA	GCAAGCGCGT	TACGCCGTGG	GTCGATGTTT	GATGTTATGG	1440
AGCAGCAACG	ATGTTACGCA	GCAGCAACGA	TGTTACGCAG	CAGGGCAGTC	GCCCTAAAAC	1500
AAAGTTAGGT	GGCTCAAGTA	TGGGCATCAT	TCGCACATGT	AGGCTCGGCC	CTGACCAAGT	1560
CAAATCCATG	CGGGCTGCTC	TTGATCTTTT	CGGTCGTGAG	TTCGGAGACG	TAGCCACCTA	1620
CTCCCAACAT	CAGCCGGACT	CCGATTACCT	CGGGAACTTG	CTCCGTAGTA	AGACATTCAT	1680
CGCGCTTGCT	GCCTTCGACC	AAGAAGCGGT	TGTTGGCGCT	CTCGCGGCTT	ACGTTCTGCC	1740
CAGGTTTGAG	CAGCCGCGTA	GTGAGATCTA	TATCTATGAT	CTCGCAGTCT	CCGGCGAGCA	1800
CCGGAGGCAG	GGCATTGCCA	CCGCGCTCAT	CAATCTCCTC	AAGCATGAGG	CCAACGCGCT	1860
TGGTGCTTAT	GTGATCTACG	TGCAAGCAGA	TTACGGTGAC	GATCCCGCAG	TGGCTCTCTA	1920
TACAAAGTTG	GGCATACGGG	AAGAAGTGAT	GCACTTTGAT	ATCGACCCAA	GTACCGCCAC	1980
CTAACAATTC	GTTCAAGCCG	AGATCGGCTT	CCCAATTGGC	CCAGCGCGTC	GATTCGGGCA	2040
TTTGCCATAT	CAATGGACCG	ACTGTGCATG	ACGAGGCTCA	GATGCCATTC	GGTGGGGTGA	2100
AGTCCAGCGG	CTACGGCAGC	TTCGGCAGTC	GAGCATCGAT	TGAGCACTTT	ACCCAGCTGC	2160

				•	
	1.				
7					
			·		
		*			G.
					•

GCTGGCTGAC CATTCAGAAT GGCCCGCGC ACTATCCAAT CTAAATCGAT CTTCGGGCGC 2220
CGCGGGCATC ATGCCCGCGG CGCTCGCCTC ATTTCAATCT CTAACTTGAT AAAAACAGAG 2280
CTGTTCTCCG GTCTTGGTGG ATCAAGGCCA GTCGCGGAGA GTCTCGAAGA GGAGAGTACA 2340
CTGAACGCCG AGTCCACATT GCAACCGCAG GCATCATCAT GCTCTGCTCA GCCACGCTAC 2400
CGCAGTGTGT CTCTCGAGGC GCTTCTTCCC TTCCCGGTG GAATTC 2460
CGGATGCGTT CTCTCGAGGC GCTTCTTCCC TTCCCGGTG GAATTC 2506

PCT/EP99/07952

ž.							-
						ě	
				ti			
	:			74 <sub>1</sub>			
							À
							· · · · · · · · · · · · · · · · · · ·
					-		

GAATTCCAAT	AATGACAATA	ATGAGGAGTG	CCCAATGTTT	CACGTGCCCC	TGCTTATTGG	60
TGGTAAGCCT	TGTTCAGCAT	CTGATGAGCG	CACCTTCGAG	CGTCGTAGCC	CGCTGACCGG	120
AGAAGTGGTA	TCGCGCGTCG	CTGCTGCCAG	TTTGGAAGAT	GCGGACGCCG	CAGTGGCCGC	180
TGCACAGGCT	GCGTTTCCTG	AATGGGCGGC	GCTTGCTCCG	AGCGAACGCC	GTGCCCGACT	240
GCTGCGAGCG	GCGGATCTTC	TAGAGGACCG	TTCTTCCGAG	TTCACCGCCG	CAGCGAGTGA	300
AACTGGCGCA	GCGGGAAACT	GGTATGGGTT	TAACGTTTAC	CTGGCGGCGG	GCATGTTGCG	360
GGAAGCCGCG	GCCATGACCA	CACAGATTCA	GGGCGATGTC	ATTCCGTCCA	ATGTGCCCGG	420
TAGCTTTGCC	ATGGCGGTTC	GACAGCCATG	TGGCGTGGTG	CTCGGTATTG	CGCCTTGGAA	480
TGCTCCGGTA	ATCCTTGGCG	TACGGGCTGT	TGCGATGCCG	TTGGCATGCG	GCAATACCGT	540
GGTGTTGAAA	AGCTCTGAGC	TGAGTCCCTT	TACCCATCGC	CTGATTGGTC	AGGTGTTGCA	600
TGATGCTGGT	CTGGGGGATG	GCGTGGTGAA	TGTCATCAGC	AATGCCCCGC	AAGACGCTCC	660
TGCGGTGGTG	GAGCGACTGA	TTGCAAATCC	TGCGGTACGT	CGAGTGAACT	TCACCGGTTC	720
GACCCACGTT	GGACGGATCA	TTGGTGAGCT	GTCTGCGCGT	CATCTGAAGC	CTGCTGTGCT	780
GGAATTAGGT	GGTAAGGCTC	CGTTCTTGGT	CTTGGACGAT	GCCGACCTCG	ATGCGGCGGT	840
CGAAGCGGCG	GCCTTTGGTG	CCTACTTCAA	TCAGGGTCAA	ATCTGCATGT	CCACTGAGCG	900
TCTGATTGTG	ACAGCAGTCG	CAGACGCCTT	TGTTGAAAAG	CTGGCGAGGA	AGGTCGCCAC	960
ACTGCGTGCT	GGCGATCCTA	ATGATCCGCA	ATCGGTCTTG	GGTTCGTTGA	TTGATGCCAA	1020
TGCAGGTCAA	CGCATCCAGG	TTCTGGTCGA	TGATGCGCTC	GCAAAAGGCG	CGCAATGGAA	1080
TTGGCCCAGC	GCGTCGATTC	GGGCATTTGC	CATATCAATG	GACCGACTGT	GCATGACGAG	1140
GCTCAGATGC	CATTCGGTGG	GGTGAAGTCC	AGCGGCTACG	GCAGCTTCGG	CAGTCGAGCA	1200
TCGATTGAGC	ACTTTACCCA	GCTGCGCTGG	CTGACCATTC	AGAATGGCCC	GCGGCACTAT	1260
CCAATCTAAA	TCGATCTTCG	GGCGCCGCGG	GCATCATGCC	CGCGGCGCTC	GCCTCATTTC	1320
AATCTCTAAC	TTGATAAAAA	CAGAGCTGTT	CTCCGGTCTT	GGTGGATCAA	GGCCAGTCGC	1380
GGAGAGTCTC	GAAGAGGAGA	GTACAGTGAA	CGCCGAGTCC	ACATTGCAAC	CGCAGGCATC	1440
ATCATGCTCT	GCTCAGCCAC	GCTACCGCAG	TGTGTCGATT	GGTCATCCTC	CGGTTGAGGT	1500
TACGCAAGAC	GCTGGAGGTA	TTGTCCGGAT	GCGTTCTCTC	GAGGCGCTTC	TTCCCTTCCC	1560
GGGTGGAATT	С					1571

		a
		,
4		
i de la companya de		

	GTCGGCGAAA					60
GTGACGAGGC	CACACTGTGA	GTTGGTCAGG	GGGGGCTTAC	TCGGCGTTTT	CCGACACTGC	120
	GCAGTGCGCA			GGTGCCCTGT		180
	TTAGGGGTAA					240
	GATAGCGTAC					300
	CACATTGGCT					360
GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	GCTTCGGCAG	GCCGGCGAGC	AGATTTCCCA	42.0
11000001011	CACGTGCTGT					480
GTATACACAC	CGGGGCGGGT	TCCGCCTCGG	TGCGCCCGTT	GAGTTCAAGG	ATTTTTTGTG	540
GGAGGCATTG	TTTGATCCTG	CTCCAGGACT	CGACATGATC	GCTACCGCAG	AAAACCTGGG	600
GACAGCAAGC	GAACCGGAAT	TGCCAGCTGG	GGCGCCCTCT	GGTAAGGTTG	GGAAGCCCTG	660
CAAAGTAAAC	TGGATGGCTT	TCTTGCCGCC	AAGGATCTGA	TGGCGCAGGG	GATCAAGATC	720
TGATCAAGAG	ACAGGATGAG	GATCGTTTCG	CATGATTGAA	CAAGATGGAT	TGCACGCAGG	780
TTCTCCGGCC	GCTTGGGTGG	AGAGGCTATT	CGGCTATGAC	TGGGCACAAC	AGACAATCGG	840
CTGCTCTGAT	GCCGCCGTGT	TCCGGCTGTC	AGCGCAGGGG	CGCCCGGTTC	TTTTTGTCAA	900
GACCGACCTG	TCCGGTGCCC	TGAATGAACT	GCAGGACGAG	GCAGCGCGGC	TATCGTGGCT	960
GGCCACGACG	GGCGTTCCTT	GCGCAGCTGT	GCTCGACGTT	GTCACTGAAG	CGGGAAGGGA	1020
CTGGCTGCTA	TTGGGCGAAG	TGCCGGGGCA	GGATCTCCTG	TCATCTCACC	TTGCTCCTGC	1080
CGAGAAAGTA	TCCATCATGG	CTGATGCAAT	GCGGCGGCTG	CATACGCTTG	ATCCGGCTAC	1140
	GACCACCAAG					1200
CGGTCTTGTC	GATCAGGATG	ATCTGGACGA	AGAGCATCAG	GGGCTCGCGC	CAGCCGAACT	1260
GTTCGCCAGG	CTCAAGGCGC	GCATGCCCGA	CGGCGAGGAT	CTCGTCGTGA	CCCATGGCGA	1320
	CCGAATATCA			TCTGGATTCA		1380
	GTGGCGGACC					1440
	GGCGAATGGG					1500
	ATCGCCTTCT					1560
	CCGACCAAGC					1620
	GCTGTTACGG				GTCGAGCAAT	1680
	CGGAAGGCAA				TAGCAGTCTT	1740
	TCCCGATTGA					1800
	GTGGACGGCG					1860
	GCTAGGATAC					1920
GCTCGGCCCT	GCGCCCGCGA				GTTTGAGGGA	1980
TATCGACCTC				GTTCTAGCGG	TACAGCATGA	2040
	GAGCACTCAA					2100
GCTTGCCGCG	ACCGGATTGC	GTCTCTGCAT	GACCCTCGCT	CACCAATTGC	AAGCTAATAA	2160

		• -•
		<u>-</u>
	•	
		•
	•	
		•
		•
•		

CTTTCGATAT	GGAATTGCCT	CGGCATGCAT	TGGTGGGGGA	CAGGGGATGG	CGGTTCTTTT	2220
AGAGAATCCC	CACTTCGGTT	CGTCCTCTGC	ACGAAGTTCG	ATGATTAACA	GAGTTGACCA	2280
CTATCCACTG	AGCTAACGGG	CATCTCCTTT	GTTGCTTTGA	GGTGGCGCAC	GAAGGAGGC	2340
TCGAAAATCT	CTGCTAAAAA	CAAGAAGAAG	GAACAGGGAA	CATGATTAGT	TTCGCTCGTA	2400
TGGCAGAAAG	TTTAGGAGTC	CAGGCTAAAC	TTGCCCTTGC	CTTCGCACTC	GTATTATGTG	2460
TCGGGCTGAT	TGTTACCGGC	ACGGGTTTCT	ACAGTGTACA	TACCTTGTCA	GGGTTGGTGG	2520
GAATTC						2526

				\$ "
				1
		·		
4.1				
				-47
	*			•
				· ·
			-	

	GTCGGCGAAA					60
GTGACGAGGC	CACACTGTGA	GTTGGTCAGG	GGGGGCTTAC	TCGGCGTTTT	CCGACACTGC	120
	GCAGTGCGCA		TGATTGCGGG			180
GCCTATCGAC	TTAGGGGTAA	AGGTCGCTCG	CGAAGTTCTG	ATGCGTGCGT	CGCTTGAACC	240
ACAAATGGTC	GATAGCGTAC	TCGCAGGCTC	TATGGCTCAA	GCAAGCTTTG	ATGCTTACCT	300
GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	TGTTCCCAAG	TCGGTTCCGG	CCTTGGGGGT	360
GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	GCTTCGGCAG	GCCGGCGAGC	AGATTTCCCA	420
AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGC	AGAGTCCATG	TCGCGTAACC	CCATCGCGTC	480
GTATACACAC	CGGGGCGGGT	TCCGCCTCGG	TGCGCCCGTT	GAGTTCAAGG	ATTTTTTGTG	540
GGAGGCATTG	TTTGATCCTG	CTCCAGGACT	CGACATGATC	GCTACCGCAG	AAAACCTGGG	600
GGAGAGGCGG	TTTGCGTATT	GGGCGCATGC	ATAAAAACTG	TTGTAATTCA	TTAAGCATTC	660
TGCCGACATG	GAAGCCATCA	CAAACGGCAT	GATGAACCTG	AATCGCCAGC	GGCATCAGCA	720
CCTTGTCGCC	TTGCGTATAA	TATTTGCCCA	TGGACGCACA	CCGTGGAAAC	GGATGAAGGC	780
ACGAACCCAG	TTGACATAAG	CCTGTTCGGT	TCGTAAACTG	TAATGCAAGT	AGCGTATGCG	840
CTCACGCAAC	TGGTCCAGAA	CCTTGACCGA	ACGCAGCGGT	GGTAACGGCG	CAGTGGCGGT	900
TTTCATGGCT	TGTTATGACT	GTTTTTTGT	ACAGTCTATG	CCTCGGGCAT	CCAAGCAGCA	960
AGCGCGTTAC	GCCGTGGGTC	GATGTTTGAT	GTTATGGAGC	AGCAACGATG	TTACGCAGCA	1020
GCAACGATGT	TACGCAGCAG	GGCAGTCGCC	CTAAAACAAA	GTTAGGTGGC	TCAAGTATGG	1080
GCATCATTCG	CACATGTAGG	CTCGGCCCTG	ACCAAGTCAA	ATCCATGCGG	GCTGCTCTTG	1140
ATCTTTTCGG	TCGTGAGTTC	GGAGACGTAG	CCACCTACTC	CCAACATCAG	CCGGACTCCG	1200
ATTACCTCGG	GAACTTGCTC	CGTAGTAAGA	CATTCATCGC	GCTTGCTGCC	TTCGACCAAG	1260
AAGCGGTTGT	TGGCGCTCTC	GCGGCTTACG	TTCTGCCCAG	GTTTGAGCAG	CCGCGTAGTG	1320
	CTATGATCTC					1380
CGCTCATCAA	TCTCCTCAAG	CATGAGGCCA	ACGCGCTTGG	TGCTTATGTG	ATCTACGTGC	1440
AAGCAGATTA	CGGTGACGAT	CCCGCAGTGG	CTCTCTATAC	AAAGTTGGGC	ATACGGGAAG	1500
	CTTTGATATC					1560
	ATTGAGGGCG					1620
	GTTCGATTTA					1680
	GATCGTGACA					1740
TGAAGCCTGT	TCATTCTGGC	GGGGTGCAGA	CTGCGGGCAA	CAGCTGTGCC	GTAGTGGACG	1800
	GGCTTTGGTG				TTGGCTAGGA	1860
	CTCCGTAGTC					1920
	GCTGCTTGCG					1980
	GCAGGCCGCC					2040
	TATTTGGGGC					2100
TGCGTCTCTG	CATGACCCTC	GCTCACCAAT	TGCAAGCTAA	TAACTTTCGA	TATGGAATTG	2160

			•
5			
			<i>,</i>
	3		•/

CCTCGGCATG	CATTGGTGGG	GGACAGGGGA	TGGCGGTTCT	TTTAGAGAAT	CCCCACTTCG	2220
GTTCGTCCTC	TGCACGAAGT	TCGATGATTA	ACAGAGTTGA	CCACTATCCA	CTGAGCTAAC	2280
GGGCATCTCC	TTTGTTGCTT	TGAGGTGGCG	CACGAAGGAG	GGCTCGAAAA	TCTCTGCTAA	2340
AAACAAGAAG	AAGGAACAGG	GAACATGATT	AGTTTCGCTC	GTATGGCAGA	AAGTTTAGGA	2400
GTCCAGGCTA	AACTTGCCCT	TGCCTTCGCA	CTCGTATTAT	GTGTCGGGCT	GATTGTTACC	2460
GGCACGGGTT	TCTACAGTGT	ACATACCTTG	TCAGGGTTGG	TGGGAATTC		2509

\*

•

GAATTCCGCG	GTCGGCGAAA	GTTGATGCGC	TGTATCGTGG	TGAAGATCAA	TCCATGCTGC	60
GTGACGAGGC	CACACTGTGA	GTTGGTCAGG	GGGGGCTTAC	TCGGCGTTTT	CCGACACTGC	120
	GCAGTGCGCA	CCCCCTGGAT	TGATTGCGGG	GGTGCCCTGT	CGCTGGTGTC	180
GCCTATCGAC	TTAGGGGTAA	AGGTCGCTCG	CGAAGTTCTG	ATGCGTGCGT	CGCTTGAACC	240
ACAAATGGTC	GATAGCGTAC	TCGCAGGCTC	TATGGCTCAA		ATGCTTACCT	300
GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	TGTTCCCAAG	TCGGTTCCGG		360
GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	GCTTCGGCAG			420
AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGC	AGAGTCCATG	TCGCGTAACC		480
GTATACACAC	CGGGGCGGGT		TGCGCCCGTT	GAGTTCAAGG		540
GGAGGCATTG	TTTGATCCTG		CGACATGATC			600
GCGCATTGAG	GGCGCAAGAG		TTGACCAAGA			660
AACAGTTCGA	TTTAGAGGGC	TACAACAGTC	GAGCAATTGA	ACTGCCTCGG	AAGGCAAAAT	720
TGTTGATCGT			CAGTCTTTGA		CGATTGAAGC	780
CTGTTCATTC	TGGCGGGGTG	CAGACTGCGG	GCAACAGCTG	TGCCGTAGTG	GACGGCGCCG	840
CGGCGGCTTT			CGACACAGCC		AGGATACTGG	900
CTACCTCCGT	AGTCGGGATC	GAGCCCGAGC	ATATGGGGCT		CCCGCGATTC	960
GCCTGCTGCT	TGCGCGTAGT	GATCTTAGTT	TGAGGGATAT		GAGATAAACG	1020
AGGCGCAGGC	CGCCCAAGTT	CTAGCGGTAC	AGCATGAATT		CACTCAAAAC	1080
TTAATATTTG	GGGCGGGCC	ATTGCACTTG			GGATTGCGTC	1140
TCTGCATGAC	CCTCGCTCAC	CAATTGCAAG	CTAATAACTT			1200
CATGCATTGG	TGGGGGACAG	GGGATGGCGG	TTCTTTTAGA		TTCGGTTCGT	1260
CCTCTGCACG	AAGTTCGATG	ATTAACAGAG	TTGACCACTA			1320
CTCCTTTGTT	GCTTTGAGGT	GGCGCACGAA		AAAATCTCTG		1380
GAAGAAGGAA	CAGGGAACAT					1440
GCTAAACTTG					TACCGGCACG	1500
GGTTTCTACA	GTGTACATAC	CTTGTCAGGG	TTGGTGGGAA	TTC		1543

. S

#### INTERNATIONAL SEARCH REPORT

Im Itional Application No PCT/EP 99/07952

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C12N15/00 C12N C12N1/00 C12N9/02 C12N9/00 C12N9/88 C12P7/22 C12P7/24 C12P7/42 C12N9/10 According to International Patent Classification (IPC) or to both national classification and IPC Minimum documentation searched (classification system followed by classification symbols) C12N C12P IPC 7 Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practical, search terms used) C. DOCUMENTS CONSIDERED TO BE RELEVANT Citation of document, with indication, where appropriate, of the relevant passages Relevant to claim No. X EP 0 845 532 A (HAARMANN & REIMER GMBH) 1,3-6,3 June 1998 (1998-06-03) 11-16 cited in the application the whole document WO 97 35999 A (INST OF FOOD RESEARCH χ 1,4,5, ; NARBAD ARJAN (GB); RHODES MICHAEL JOHN 12-15 CHAR) 2 October 1997 (1997-10-02) cited in the application the whole document -/--Further documents are listed in the continuation of box C. X Patent family members are listed in annex. X Special categories of cited documents: later document published after the international filling date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the "A" document defining the general state of the art which is not considered to be of particular relevance invention "E" earlier document but published on or after the international "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such docu-ments, such combination being obvious to a person skilled "O" document referring to an oral disclosure, use, exhibition or other means document published prior to the international filing date but later than the priority date claimed in the art. "&" document member of the same patent family Date of the actual completion of the international search Date of mailing of the international search report 7 June 2000 26/06/2000 Name and mailing address of the ISA Authorized officer European Patent Office. P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, 7x, 31 651 epo ni. Fax: (+31-70) 340-3016 Oderwald, H

## INTERNATIONAL SEARCH REPORT

Int Ational Application No PCT/EP 99/07952

	ation) DOCUMENTS CONSIDERED TO BE RELEVANT	
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
x	GASSON M J ET AL.: "Metabolism of ferulic acid to vanillin" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 273, no. 7, 13 February 1998 (1998-02-13), pages 4163-4170, XP002139585 cited in the application the whole document	1,4,5, 12-15
X	DATABASE WPI Section Ch, Week 199340 Derwent Publications Ltd., London, GB; Class B05, AN 1993-316614 XP002139586 & JP 05 227980 A (TAKASAGO PERFUMERY CO LTD), 7 September 1993 (1993-09-07) abstract	1,4,5, 12,15
A	EP 0 583 687 A (HAARMANN & REIMER GMBH) 23 February 1994 (1994-02-23) cited in the application the whole document	
	:	

# INTERNATIONAL SEARCH REPORT

Information on patent family members

Inc. Itional Application No PCT/EP 99/07952

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
EP 0845532	A	03-06-1998	DE 19649655 A JP 10155496 A	04-06-1998 16-06-1998
WO 9735999	A	02-10-1997	AU 2038597 A CA 2250043 A EP 0904396 A	17-10-1997 02-10-1997 31-03-1999
JP 5227980	Α	07-09-1993	NONE	
EP 0583687	A	23-02-1994	DE 4227076 A DE 59309293 D JP 6153924 A US 5371013 A US 5510252 A US 5976863 A	24-02-1994 25-02-1999 03-06-1994 06-12-1994 23-04-1996 02-11-1999

1. 

----

## INTERNATIONALER RECHERCHENBERICHT

PCT/EP 99/07952

A. KLASSI IPK 7	Fizierung des anmeldungsgegenstandes C12N15/00 C12N1/00 C12N9/0 C12N9/10 C12P7/22 C12P7/2		C12N9/88							
Nach der Int	ternationalen Patentklassifikation (IPK) oder nach der nationalen Kl	assifikation und der IPK								
B. RECHERCHIERTE GEBIETE										
Recherchier IPK 7	ter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymt C12N C12P	bole )								
Recherchiei	te aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, s	soweit diese unter die recherchierte	n Gebiete fallen							
	r internationalen Recherche konsultierte elektronische Datenbank (	Name der Datenbank und evtl. ver	wendete Suchbegriffe)							
C. ALS WE	SENTLICH ANGESEHENE UNTERLAGEN									
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angai	be der in Betracht kommenden Teil	e Betr. Anspruch Nr.							
х	EP 0 845 532 A (HAARMANN & REIME 3. Juni 1998 (1998-06-03) in der Anmeldung erwähnt das ganze Dokument	1,3-6, 11-16								
X	WO 97 35999 A (INST OF FOOD RESE; NARBAD ARJAN (GB); RHODES MICHA CHAR) 2. Oktober 1997 (1997-10-0 in der Anmeldung erwähnt das ganze Dokument	1,4,5, 12-15								
		-/								
	ehrmen	X Siehe Anhang Patentfam	nilie							
"Besondern "A" Veröffer aber n "E" ätteres Anmel "L" Veröffer scheir andern soll oc ausge "O" Veröffe enne B "P" Veröffer dem b	*Besondere Kategorien von angegebenen Veröffentlichungen:  *A* Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist steres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist und mit der Anmeldedatum veröffentlichtung internationalen Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist "X* Veröffentlichung, von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden "Y* Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Aussteltung oder andere Maßnahmen bezieht dem beanspruchten Prioritätsdatum veröffentlicht worden ist und mit der Anmeldedatum veröffentlichtung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung, die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung van ausgeführt.  "X* Veröffentlichung, die sich auf erimerischer Tätigkeit beruhend betrachtet werden veröffentlichung, die sich auf erimerischer Tätigkeit beruhend betrachtet werde									
	. Juni 2000	26/06/2000	· mass. · · 1001 101 101 101 101 101 113							
Name und f	Postanschrift der Internationalen Recherchenbehorde Europaisches Patentamt, P.B. 5818 Patentiaan,2 NL = 2280 HV Rijswijk Tel (+31=70) 340=2040, Tx. 31 651 epo ni,	Bevollmachtigter Bedienstete Oderwald H	-							

## INTERNATIONALER RECHERCHENBERICHT

In: ationales Aktenzeichen PCT/EP 99/07952

		C1/EP 99/0/952
C.(Fortset	ung) ALS WESENTLICH ANGESEHENE UNTERLAGEN	
Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommend	en Teile Betr. Anspruch Nr.
X	GASSON M J ET AL.: "Metabolism of ferulic acid to vanillin" JOURNAL OF BIOLOGICAL CHEMISTRY, Bd. 273, Nr. 7, 13. Februar 1998 (1998-02-13), Seiten 4163-4170, XP002139585 in der Anmeldung erwähnt das ganze Dokument	1,4,5, 12-15
x	DATABASE WPI Section Ch, Week 199340 Derwent Publications Ltd., London, GB; Class B05, AN 1993-316614 XP002139586 JP 05 227980 A (TAKASAGO PERFUMERY CO LTD), 7. September 1993 (1993-09-07) Zusammenfassung	1,4,5, 12,15
A	EP 0 583 687 A (HAARMANN & REIMER GMBH) 23. Februar 1994 (1994-02-23) in der Anmeldung erwähnt das ganze Dokument	

# INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffemtlichungen, die zur selben Patentlamitie gehören

In. .tionales Aktenzeichen
PCT/EP 99/07952

	lecherchenberich Intes Patentdoku		Datum der Veröffentlichung		litglied(er) der Patentfamilie	Datum der Veröffentlichung
EP	0845532	A	03-06-1998	DE JP	19649655 A 10155496 A	04-06-1998 16-06-1998
WO	9735999	Α	02-10-1997	AU CA EP	2038597 A 2250043 A 0904396 A	17-10-1997 02-10-1997 31-03-1999
JP	5227980	Α	07-09-1993	KEIN	VE	
EP	0583687	A	23-02-1994	DE DE JP US US US	4227076 A 59309293 D 6153924 A 5371013 A 5510252 A 5976863 A	24-02-1994 25-02-1999 03-06-1994 06-12-1994 23-04-1996 02-11-1999

-



# PCT LTORGANISATION FÜR GEISTIGES EIGENTUM Internationales Büro INTERNATIONALE ANMELDUNG VERÖFFENTLICHT NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT)

(51) Internationale Patentklassifikation 7:

C12N 15/00

A2

- WO 00/26355 (11) Internationale Veröffentlichungsnummer:
- (43) Internationales Veröffentlichungsdatum:

11. Mai 2000 (11.05.00)

(21) Internationales Aktenzeichen:

PCT/EP99/07952

- (22) Internationales Anmeldedatum: 20. Oktober 1999 (20.10.99)
- (30) Prioritätsdaten:

198 50 242.7

DF. 31. Oktober 1998 (31.10.98)

- (71) Anmelder (für alle Bestimmungsstaaten ausser US): HAAR-MANN & REIMER GMBH [DE/DE]; D-37601 Holzminden (DE).
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): RABENHORST, Jürgen Weserblick 13, D-37671 Höxter (DE). STEINBÜCHEL, Alexander [DE/DE]; Rönnenthal 27, D-48341 Altenberge (DE). PRIEFERT, Horst [DE/DE]; Potthoffskamp 5, D-48291 Telgte (DE). OVERHAGE, Jörg [DE/DE]; Rotherbachstrasse 42 a, D-59132 Bergkamen (DE).
- (74) Anwalt: MANN, Volker; Bayer Aktiengesellschaft, D-51368 Leverkusen (DE).

(81) Bestimmungsstaaten: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO Patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

#### Veröffentlicht

Ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts.

- (54) Title: CONSTRUCTION OF PRODUCTION STRAINS FOR PRODUCING SUBSTITUTED PHENOLS BY SPECIFICALLY INACTIVATING GENES OF THE EUGENOL AND FERULIC ACID CATABOLISM
- (54) Bezeichnung: KONSTRUKTION VON PRODUKTIONSSTÄMMEN FÜR DIE HERSTELLUNG VON SUBSTITUIERTEN PHENOLEN DURCH GEZIELTE INAKTIVIERUNGEN VON GENEN DES EUGENOL- UND FERU-LASÄURE-KATABOLISMUS

#### (57) Abstract

The invention relates to a transformed and/or mutagenated unicellular or multicellular organism which is characterized in that enzymes of the eugenol and/or ferulic acid catabolism are deactivated in such a manner that the intermediates coniferyl alcohol, coniferyl aldehyde, ferulic acid, vanillin and/or vanillinic acid are accumulated.

#### (57) Zusammenfassung

Die vorliegende Erfindung betrifft einen transformierten und/oder mutagenisierten ein- oder mehrzelligen Organismus, der dadurch gekennzeichnet ist, daß Enzyme des Eugenol- und/oder Ferulasäure-Katabolismus derart inaktiviert sind, daß eine Akkumulation der Intermediate Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure erfolgt.

#### LEDIGLICH ZUR INFORMATION

Codes zur Identifizierung von PCT-Vertragsstaaten auf den Kopfbögen der Schriften, die internationale Anmeldungen gemäss dem PCT veröffentlichen.

AL	Albanien	ES	Spanien	LS	Lesotho	SI	Slowenien
AM	Armenien	FI	Finnland	LT	Litauen	SK	Slowakei
AT	Österreich	FR	Frankreich	LU	Luxemburg	SN	Senegal
AU	Australien	GA	Gabun	LV	Lettland	SZ	Swasiland
AZ	Aserbaidschan	GB	Vereinigtes Königreich	MC	Monaco	TD	Tschad
BA	Bosnien-Herzegowina	GE	Georgien	MD	Republik Moldau	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagaskar	T.J	Tadschikistan
BE	Belgien	GN	Guinea	MK	Die ehemalige jugoslawische	TM	Turkmenistan
BF	Burkina Faso	GR	Griechenland		Republik Mazedonien	TR	Türkei
BG	Bulgarien	HU	Ungarn	ML	Mali	TT	Trinidad und Tobago
BJ	Benin	ΙE	Irland	MN	Mongolei	UA	Ukraine
BR	Brasilien	IL	Israel	MR	Mauretanien	UG	Uganda
BY	Belarus	IS	Island	MW	Malawi	US	Vereinigte Staaten von
CA	Kanada	IT	Italien	MX	Mexiko	05	Amerika
CF	Zentralafrikanische Republik	JP	Japan	NE	Niger	UZ	Usbekistan
CG	Kongo	KE	Kenia	NL	Niederlande	VN	Vietnam
CH	Schweiz	KG	Kirgisistan	NO	Norwegen	YU	Jugoslawien
CI	Côte d'Ivoire	KP	Demokratische Volksrepublik	NZ	Neusceland	zw	Zimbabwe
CM	Kamerun		Korea	PL	Polen	2311	Zimbaowe
CN	China	KR	Republik Korea	PT	Portugal		
CU	Kuba	ΚZ	Kasachstan	RO	Rumänien		•
CZ	Tschechische Republik	LC	St. Lucia	RU	Russische Föderation		
DE	Deutschland	LI	Liechtenstein	SD	Sudan		
DK	Dänemark	LK	Sri Lanka	SE	Schweden		
EE	Estland	LR	Liberia	SG	Singapur		

Konstruktion von Produktionsstämmen für die Herstellung von substituierten Phenolen durch gezielte Inaktivierungen von Genen des Eugenol- und Ferulasäure-Katabolismus

5

Die vorliegende Erfindung betrifft die Konstruktion von Produktionsstämmen und ein Verfahren für die Herstellung substituierter Methoxyphenole, insbesondere Vanillin.

10

Die DE-A 4 227 076 (Verfahren zur Herstellung substituierter Methoxyphenole und dafür geeigneter Mikroorganismus) beschreibt die Herstellung substituierter Methoxyphenole mit einer neuen *Pseudomonas* sp.. Ausgangsmaterial ist hier Eugenol und die Produkte sind Ferulasäure, Vanillinsäure, Coniferylalkohol und Coniferylaldehyd.

15

Ebenfalls 1995 erscheint ein umfangreiches Review über die Biotransformationsmöglichkeiten mit Ferulasäure von Rosazza et al. (Biocatalytic transformation of ferulic acid: an abundant aromatic natural product; J. Ind. Microbiol. **15:**457-471).

20

Die Gene und Enzyme zur Synthese von Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und Vanillinsäure aus *Pseudomonas* sp. wurden in EP-A 0 845 532 beschrieben.

25

30

Die Enzyme zur Umsetzung von *trans*-Ferulasäure zu *trans*-Feruloyl-SCoA Ester und weiter zum Vanillin, sowie das Gen für die Spaltung des Esters wurden vom Institute of Food Research, Norwich, GB, in WO 97/35999 beschrieben. 1998 erscheint der Inhalt des Patents auch als wissenschaftliche Publikationen (Gasson et al. 1998. Metabolism of ferulic acid to vanillin. J. Biol. Chem. 273:4163-4170; Narbad and Gasson 1998. Metabolism of ferulic acid via vanillin using a novel CoAdependent pathway in a newly isolated strain of *Pseudomonas fluorescens*. Microbiology 144:1397 - 1405).

Die DE-A 195 32 317 beschreibt die fermentative Gewinnung von Vanillin aus Ferulasäure mit *Amycolatopsis* sp. in hohen Ausbeuten.

Die bekannten Verfahren haben den Nachteil, daß entweder nur sehr geringe Ausbeuten an Vanillin erzielt werden, oder von relativ teuren Edukten ausgegangen wird. Bei dem letztgenannten Verfahren (DE-A 195 32 317) werden zwar hohe Ausbeuten erzielt, jedoch bedingt der Einsatz von *Pseudomonas* sp. HR199 und *Amycolatopsis* sp. HR167 für die Biotransformation von Eugenol zu Vanillin eine zweistufige Fermentationsführung und somit einen erheblichen Kosten- und Zeitaufwand.

Aufgabe der vorliegenden Erfindung ist es daher, Organismen zu konstruieren, die in der Lage sind den preiswerten Rohstoff Eugenol in einem einstufigen Prozeß zu Vanillin umzusetzen.

15

20

25

30

5

10

Diese Aufgabe wird durch die Konstruktion von Produktionsstämmen ein- oder mehrzelliger Organismen gelöst, die dadurch gekennzeichnet sind, daß Enzyme des Eugenol- und/oder Ferulasäure-Katabolismus derart inaktiviert sind, daß eine Akkumulation der Intermediate Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure erfolgt.

Der Produktionsstamm kann einzellig oder mehrzellig sein. Demgemäß können Gegenstand der Erfindung Mikroorganismen, Pflanzen oder Tiere sein. Darüber hinaus können auch Extrakte die aus dem Produktionsstamm gewonnen werden zum Einsatz kommen. Erfindungsgemäß werden vorzugsweise einzellige Organismen eingesetzt. Hierbei kann es sich um Mikroorganismen, tierische oder pflanzliche Zellen handeln. Besonders bevorzugt ist erfindungsgemäß der Einsatz von Pilzen und Bakterien. Höchst bevorzugt sind Bakterienarten. Unter den Bakterien können insbesondere Rhodococcus-, Pseudomonas- und Escherichia-Arten nach Veränderung des Eugenol- und/oder Ferulasäure-Katabolismus zum Einsatz kommen.

Die Gewinnung der erfindungsgemäß einsetzbaren Organismen kann im einfachsten Fall mittels bekannter, konventioneller mikrobiologischer Methoden erfolgen. So kann die Enzymaktivität der am Eugenol- und/oder Ferulasäure-Katabolismus beteiligten Proteine durch den Einsatz von Enzym-Hemmstoffen verändert werden. Darüber hinaus kann die Enzymaktivität der am Eugenol- und/oder Ferulasäure-Katabolismus beteiligten Proteine durch Mutation der für diese Proteine kodierenden Gene verändert werden. Derartige Mutationen können nach klassischen Methoden ungerichtet erzeugt werden, wie beispielsweise durch UV-Bestrahlung oder mutationsauslösende Chemikalien.

10

15

5

Ebenso sind gentechnische Methoden zur Gewinnung der erfindungsgemäßen Organismen geeignet, wie Deletionen, Insertionen und/oder Nukleotid-Austausche. So können beispielsweise die Gene der Organismen mit Hilfe von anderen DNA-Elementen (Ω-Elemente) inaktiviert werden. Ebenso können mittels geeigneter Vektoren Austausche der intakten Gene gegen veränderte und/oder inaktivierte Gen-Strukturen durchgeführt werden. Die zu inaktivierenden Gene und die für die Inaktivierung eingesetzten DNA-Elemente können dabei durch klassische Klonierungstechniken oder durch Polymerase-Kettenreaktionen (PCR) gewonnen werden.

In einer möglichen Ausgestaltung der Erfindung kann beispielsweise der Eugenol-20 sowie der Ferulasäure-Katabolismus durch Ω-Element-Insertion oder Einführen von Deletionen in entsprechende Gene verändert werden. Hierbei können die Funktionen der Gene, die für Dehydrogenasen, Synthetasen, Hydratasen-Aldolasen, Thiolasen oder Demethylasen kodieren, mittels der oben genannten gentechnischen Methoden 25 inaktiviert werden, so daß die Erzeugung der betreffenden Enzyme blockiert ist. Vorzugsweise handelt es sich um die Gene, die für Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA-Synthetasen, Enoyl-CoA-Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen kodieren. Ganz besonders bevorzugt sind Gene, die die in der 30 EP-A 0845532 angegebenen Aminosäuresequenzen kodieren und/oder deren Allelvariationen kodierenden Nukleotidsequenzen.

Gegenstand der Erfindung sind demgemäß auch Gen-Strukturen zur Herstellung transformierter Organismen und Mutanten.

5

10

15

20

25

30

Vorzugsweise werden Gen-Strukturen zur Gewinnung der Organismen und Mutanten eingesetzt, bei denen die für Dehydrogenasen, Synthetasen, Hydratasen-Aldolasen, Thiolasen oder Demethylasen kodierenden Nukleotidsequenzen inaktiviert sind. Besonders bevorzugt sind Gen-Strukturen, bei dene die für Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA-Synthetasen, Enoyl-CoA-Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen kodierenden Nukleotidsequenzen inaktiviert sind. Ganz besonders bevorzugt sind Gen-Strukturen, die die in den Figuren 1a bis 1r angegebenen Strukturen mit den in den Figuren 2a bis 2r wiedergegebenen Nukleotidsequenzen und/oder deren Allelvariationen kodierenden Nukleotidsequenzen von 1 bis 18.

Die Erfindung schließt auch die Teilsequenzen dieser Gen-Strukturen sowie funktionelle Äquivalente ein. Unter funktionellen Äquivalenten sind solche Derivate der DNA zu verstehen, bei denen einzelne Nukleobasen ausgetauscht worden sind (Wobbelaustausche), ohne die Funktion zu ändern. Auch auf Proteinebene können Aminosäuren ausgetauscht werden, ohne daß eine Veränderung der Funktion die Folge ist.

Den Gen-Strukturen können ein oder mehrere DNA-Sequenzen vor- und/oder nachgeschaltet sein. Durch Klonierung der Gen-Strukturen sind Plasmide bzw. Vektoren erhältlich, die zur Transformation und/oder Transfektion eines Organismus und/oder zur konjugativen Übertragung in einen Organismus geeignet sind.

Gegenstand der Erfindung sind ferner Plasmide und/oder Vektoren zur Herstellung der erfindungsgemäßen transformierten Organismen und Mutanten. Diese enthalten demgemäß die beschriebenen Gen-Strukturen. Die vorliegende Erfindung betrifft

WO 00/26355 PCT/EP99/07952

demgemäß auch Organismen, die die genannten Plasmide und/oder Vektoren enthalten.

Die Art der Plasmide und/oder Vektoren hängt von deren Einsatzzweck ab. Um z. B. die intakten Gene des Eugenol- und/oder Ferulasäure-Katabolismus in Pseudomonaden gegen die durch Omega-Elemente inaktivierten Gene austauschen zu können, benötigt man Vektoren, die einerseits in Pseudomonaden übertragen werden können (konjugativ übertragbare Plasmide), andererseits dort jedoch nicht repliziert werden können und somit in Pseudomonaden instabil sind (sogenannte Suizid-Plasmide). DNA-Abschnitte, die mit Hilfe eines solchen Plasmidsystems in Pseudomonaden übertragen werden, können nur erhalten bleiben, wenn sie durch homologe Rekombination in das Genom der Bakterienzelle integriert werden.

Die beschriebenen Gen-Strukturen, Vektoren und Plasmide können zur Herstellung verschiedener transformierter Organismen oder Mutanten verwendet werden. Mittels der gennanten Gen-Strukturen können intakte Nukleinsäuresequenzen gegen veränderte und/oder inaktivierte Gen-Strukturen ausgetauscht werden. In den durch Transformation oder Transfektion oder Konjugation erhältlichen Zellen erfolgt durch homologe Rekombination ein Austausch des intakten Gens gegen die veränderte und/oder inaktivierte Gen-Struktur, wodurch die resultierenden Zellen nur noch über die veränderte und/oder inaktivierte Gen-Struktur im Genom verfügen. So können erfindungsgemäß vorzugsweise Gene derart verändert und/oder inaktiviert werden, daß die betreffenden Organismen Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure zu erzeugen vermögen.

25

30

5

10

15

20

Erfindungsgemäß derart konstruierte Produktionsstämme sind beispielsweise Mutanten des Stammes *Pseudomonas* sp. HR199 (DSM 7063) der in der DE-A 4 227 076 und der EP-A 0845532 genau beschrieben wurde, wobei sich unter anderem die entsprechenden Genstrukturen aus den Figuren 1a bis 1r in Verbindung mit den Figuren 2a bis 2r ergeben:

- 1. Pseudomonas sp. HR199calAΩKm, enthaltend das durch ΩKm inaktivierte calA-Gen an Stelle des intakten calA-Gens kodierend für Coniferylalkohol-Dehydrogenase (Fig. 1a; Fig. 2a).
- Pseudomonas sp. HR199calAΩGm, enthaltend das durch ΩGm inaktivierte
   calA-Gen an Stelle des intakten calA-Gens kodierend für Coniferylalkohol-Dehydrogenase (Fig. 1b; Fig. 2b).
  - 3. Pseudomonas sp. HR199calAΔ, enthaltend das durch Deletion inaktivierte calA-Gen an Stelle des intakten calA-Gens kodierend für Coniferylalkohol-Dehydrogenase (Fig. 1c; Fig. 2c).
- Pseudomonas sp. HR199calBΩKm, enthaltend das durch ΩKm inaktivierte calB-Gen an Stelle des intakten calB-Gens kodierend für Coniferylaldehyd-Dehydrogenase (Fig. 1d; Fig. 2d)

- 5. Pseudomonas sp. HR199 $calB\Omega$ Gm, enthaltend das durch  $\Omega$ Gm inaktivierte calB-Gen an Stelle des intakten calB-Gens kodierend für Coniferylaldehyd-Dehydrogenase (Fig. 1e; Fig. 2e).
- 6. Pseudomonas sp. HR199calBΔ, enthaltend das durch Deletion inaktivierte calB-Gen an Stelle des intakten calB-Gens kodierend für Coniferylaldehyd-Dehydrogenase(Fig.1f; Fig. 2f).
- 7. Pseudomonas sp. HR199fcsΩKm, enthaltend das durch ΩKm inaktivierte fcs 20 Gen an Stelle des intakten fcs-Gens kodierend für Ferulasäure-CoA-Synthetase (Fig.1g; Fig. 2g).
  - 8. Pseudomonas sp. HR199 $fcs\Omega$ Gm, enthaltend das durch  $\Omega$ Gm inaktivierte fcs-Gen an Stelle des intakten fcs-Gens kodierend für Ferulasäure-CoA-Synthetase (Fig.1h; Fig. 2h).
- 9. Pseudomonas sp. HR199fcsΔ, enthaltend das durch Deletion inaktivierte fcs-Gen an Stelle des intakten fcs-Gens kodierend für Ferulasäure-CoA-Synthetase (Fig. 1i; Fig. 2i).
- 10. Pseudomonas sp. HR199echΩKm, enthaltend das durch ΩKm inaktivierte ech-Gen an Stelle des intakten ech-Gens kodierend für Enoyl-CoA-Hydratase-Aldolase (Fig.1j; Fig. 2j).

- 11. Pseudomonas sp. HR199echΩGm, enthaltend das durch ΩGm inaktivierte ech-Gen an Stelle des intakten ech-Gens kodierend für Enoyl-CoA-Hydratase-Aldolase (Fig.1k; Fig. 2k).
- 12. Pseudomonas sp. HR199echΔ, enthaltend das durch Deletion inaktivierte ech-Gen an Stelle des intakten ech-Gens kodierend für Enoyl-CoA-Hydratase-Aldolase (Fi.11; Fig. 21).

15

- 13. Pseudomonas sp. HR199aatΩKm, enthaltend das durch ΩKm inaktivierte aat-Gen an Stelle des intakten aat-Gens kodierend für beta-Ketothiolase (Fig. 1m; Fig. 2m).
- 10 14. Pseudomonas sp. HR199aatΩGm, enthaltend das durch ΩGm inaktivierte aat-Gen an Stelle des intakten aat-Gens kodierend für beta-Ketothiolase (Fig.1n; Fig. 2n).
  - 15. Pseudomonas sp. HR199aatΔ, enthaltend das durch Deletion inaktivierte aat-Gen an Stelle des intakten aat-Gens kodierend für beta-Ketothiolase (Fig.1o; 2o).
  - 16. Pseudomonas sp. HR199 $vdh\Omega$ Km, enthaltend das durch  $\Omega$ Km inaktivierte vdh-Gen an Stelle des intakten vdh-Gens kodierend für Vanillin-Dehydrogenase (Fig.1p; Fig. 2p).
- 17. Pseudomonas sp. HR199vdhΩGm, enthaltend das durch ΩGm inaktivierte
   20 vdh-Gen an Stelle des intakten vdh-Gens kodierend für Vanillin-Dehydrogenase (Fig.1q; Fig. 2q).
  - 18. Pseudomonas sp. HR199vdhΔ, enthaltend das durch Deletion inaktivierte vdh-Gen an Stelle des intakten vdh-Gens kodierend für Vanillin-Dehydrogenase (Fig.1r; Fig. 2r).
- 25 19. Pseudomonas sp. HR199 $vdhB\Omega$ Km, enthaltend das durch  $\Omega$ Km inaktivierte vdhB-Gen an Stelle des intakten vdhB-Gens kodierend für Vanillin-Dehydrogenase II.
  - 20. Pseudomonas sp. HR199vdhBΩGm, enthaltend das durch ΩGm inaktivierte vdhB-Gen an Stelle des intakten vdhB-Gens kodierend für Vanillin-Dehydrogenase II.

- Pseudomonas sp. HR199vdhBΔ, enthaltend das durch Deletion inaktivierte vdhB-Gen an Stelle des intakten vdhB-Gens kodierend für Vanillin-Dehydrogenase II.
- 22. Pseudomonas sp. HR199 $adh\Omega$ Km, enthaltend das durch  $\Omega$ Km inaktivierte adh-Gen an Stelle des intakten adh-Gens kodierend für Alkohol-Dehydrogenase.
- 23. Pseudomonas sp. HR199 $adh\Omega$ Gm, enthaltend das durch  $\Omega$ Gm inaktivierte adh-Gen an Stelle des intakten adh-Gens kodierend für Alkohol-Dehydrogenase.
- 24. Pseudomonas sp. HR199adhΔ enthaltend das durch Deletion inaktivierte adh Gen an Stelle des intakten adh-Gens kodierend für Alkohol-Dehydrogenase.
  - 25. Pseudomonas sp. HR199 $vanA\Omega$ Km, enthaltend das durch  $\Omega$ Km inaktivierte vanA-Gen an Stelle des intakten vanA-Gens kodierend für die  $\alpha$ -Untereinheit der Vanillinsäure-Demethylase.
- 15 26. Pseudomonas sp. HR199 $vanA\Omega$ Gm, enthaltend das durch  $\Omega$ Gm inaktivierte vanA-Gen an Stelle des intakten vanA-Gens kodierend für die  $\alpha$ -Untereinheit der Vanillinsäure-Demethylase.
  - 27. Pseudomonas sp. HR199 $vanA\Delta$ , enthaltend das durch Deletion inaktivierte vanA-Gen an Stelle des intakten vanA-Gens kodierend für die  $\alpha$ -Untereinheit der Vanillinsäure-Demethylase.
  - 28. Pseudomonas sp. HR199 $vanB\Omega$ Km, enthaltend das durch  $\Omega$ Km inaktivierte vanB-Gen an Stelle des intakten vanB-Gens kodierend für die  $\beta$ -Untereinheit der Vanillinsäure-Demethylase.
- 29. Pseudomonas sp. HR199vanBΩGm, enthaltend das durch ΩGm inaktivierte
   25 vanB-Gen an Stelle des intakten vanB-Gens kodierend für die β-Untereinheit der Vanillinsäure-Demethylase.
  - 30. *Pseudomonas* sp. HR199*vanB*Δ, enthaltend das durch Deletion inaktivierte *vanB*-Gen an Stelle des intakten *vanB*-Gens kodierend für die β-Untereinheit der Vanillinsäure-Demethylase.

5

10

15

Gegenstand der Erfindung ist außerdem ein Verfahren zur biotechnischen Herstellung von organischen Verbindungen. Insbesondere können mit diesem Verfahren Alkohole, Aldehyde und organische Säuren hergestellt werden. Vorzugsweise handelt es sich hierbei um Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und Vanillinsäure.

In dem erfindungsgemäßen Verfahren werden die oben beschriebenen Organismen eingesetzt. Zu den ganz besonders bevorzugten Organismen gehören Bakterien, insbesondere die *Pseudomonas*-Arten. Im einzelnen können die oben genannten *Pseudomonas*-Arten vorzugsweise für folgende Verfahren eingesetzt werden:

- 1. Pseudomonas sp. HR199cal $A\Omega$ Km, Pseudomonas sp. HR199cal $A\Omega$ Gm und Pseudomonas sp. HR199cal $A\Delta$  zur Herstellung von Coniferylalkohol aus Eugenol.
- Pseudomonas sp. HR199calBΩKm, Pseudomonas sp. HR199calBΩGm und Pseudomonas sp. HR199calBΔ zur Herstellung von Coniferylaldehyd aus Eugenol oder Coniferylalkohol.
- Pseudomonas sp. HR199fcsΩKm, Pseudomonas sp. HR199fcsΩGm, Pseudomonas sp. HR199fcsΔ, Pseudomonas sp. HR199echΩKm, Pseudomonas sp. HR199echΩGm und Pseudomonas sp. HR199echΔ zur Herstellung von Ferulasäure aus Eugenol oder Coniferylalkohol oder Coniferylaldehyd.
- Pseudomonas sp. HR199vdhΩKm, Pseudomonas sp. HR199vdhΩGm, Pseudomonas sp. HR199vdhΩ GmvdhBΩKm, Pseudomonas sp. HR199vdhΩGmvdhBΩKm, Pseudomonas sp. HR199vdhΩKmvdhBΩGm, Pseudomonas sp. HR199vdhΔvdhBΩGm und Pseudomonas sp. HR199vdhΔvdhBΩKm zur Herstellung von Vanillin aus Eugenol oder Coniferylalkohol oder Coniferylaldehyd oder Ferulasäure.

5. Pseudomonas sp. HR199vanAΩKm, Pseudomonas sp. HR199vanAΩGm, Pseudomonas sp. HR199vanAΔ, Pseudomonas sp. HR199vanBΩKm, Pseudomonas sp. HR199vanBΩGm und Pseudomonas sp. HR199vanBΔ zur Herstellung von Vanillinsäure aus Eugenol oder Coniferylalkohol oder Coniferylaldehyd oder Ferulasäure oder Vanillin.

Bevorzugtes Substrat ist Eugenol. Jedoch kann der Zusatz weiterer Substrate oder sogar der Austausch des Eugenol gegen ein anderes Substrat möglich sein.

10

5

Als Nährmedium für die erfindungsgemäß eingesetzten Organismen kommen synthetische, halbsynthetische oder komplexe Kulturmedien in Betracht. Diese können kohlenstoffhaltige und stickstoffhaltige Verbindungen, anorganische Salze, gegebenenfalls Spurenelemente sowie Vitamine enthalten.

15

Als kohlenstoffhaltige Verbindungen können Kohlenhydrate, Kohlenwasserstoffe oder organische Grundchemikalien in Betracht kommen. Beispiele für vorzugsweise verwendbare Verbindungen sind Zucker, Alkohole bzw. Zuckeralkohole, organische Säuren oder komplexe Gemische.

20

Als Zucker kommt vorzugsweise Glucose in Betracht. Als organische Säuren können vorzugsweise Zitronensäure oder Essigsäure zum Einsatz kommen. Zu den komplexen Gemischen zählen z. B. Malzextrakt, Hefeextrakt, Casein oder Caseinhydrolysat.

25

30

Als stickstoffhaltige Substrate kommen anorganische Verbindungen in Betracht. Beispiele hierfür sind Nitrate und Ammoniumsalze. Ebenso können organische Stickstoffquellen zum Einsatz kommen. Hierzu zählen Hefeextrakt, Sojamehl, Casein, Caseinhydrolysat und Maisquellwasser.

Zu den einsetzbaren anorganischen Salzen zählen beispielsweise Sulfate, Nitrate, Chloride, Carbonate und Phosphate. Als Metalle enthalten die genannte Salze vorzugsweise Natrium, Kalium, Magnesium, Mangan, Calcium, Zink und Eisen.

Die Temperatur für die Kultivierung liegt vorzugsweise im Bereich von 5 bis 100°C. Besonders bevorzugt ist der Bereich von 15 bis 60°C, höchst bevorzugt sind 22 bis 37°C.

Der pH-Wert des Mediums beträgt bevorzugt 2 bis 12. Besonders bevorzugt ist der Bereich von 4 bis 8.

10

15

30

Grundsätzlich können für die Durchführung des erfindungsgemäßen Verfahrens alle dem Fachmann bekannten Bioreaktoren eingesetzt werden. Vorzugsweise kommen alle für Submersverfahren geeigneten Vorrichtungen in Betracht. Das heißt, es können erfindungsgemäß Gefäße ohne oder mit mechanischer Mischeinrichtung eingesetzt werden. Zu ersteren zählen z. B. Schüttelapparaturen, Blasensäulen- oder Schlaufenreaktoren. Zu letzteren gehören vorzugsweise alle bekannten Vorrichtungen mit Rührern in beliebiger Gestaltung.

- Das erfindungsgemäße Verfahren kann kontinuierlich oder diskontinuierlich durchgeführt werden. Die Dauer der Fermentation bis zum Erreichen einer maximalen Produktmenge hängt von der speziellen Art des eingesetzten Organismus ab. Grundsätzlich liegen die Zeiten der Fermentation jedoch zwischen 2 und 200 Stunden.
- Im folgenden wird die Erfindung unter Bezugnahme auf Beispiele näher erläutert:

Von dem Eugenol verwertenden Stamm Pseudomonas sp. HR199 (DSM 7063) wurden gezielt Mutanten erzeugt, wobei spezifisch Gene des Eugenol-Katabolismus durch Insertion von Omega-Elementen oder durch Einführen von Deletionen inaktiviert wurden. Als Omega-Elemente dienten DNA-Abschnitte die für Antibiotikaresistenzen gegen Kanamycin ( $\Omega$ Km) und Gentamycin ( $\Omega$ Gm) codierten. Diese

Resistenzgene wurden ausgehend von Tn5 und dem Plasmid pBBR1MCS-5 mit Hilfe von Standardmethoden isoliert. Die Gene calA, calB, fcs, ech, aat, vdh, adh, vdhB, vanA und vanB, die für Coniferylalkohol-Dehydrogenase, Coniferylaldehyd-Dehydrogenase, Ferulasäure-CoA Synthetase, Enoyl-CoA Hydratase-Aldolase, beta-Ketothiolase, Vanillin-Dehydrogenase, Alkohol-Dehydrogenase, Vanillin-Dehydrogenase II und Vanillinsäure-Demethylase codieren wurden ausgehend von genomischer DNA des Stammes Pseudomonas sp. HR199 mit Hilfe von Standardmethoden isoliert und in pBluescript SK kloniert. Aus diesen Genen wurden durch Verdauung mit geeigneten Restriktionsendonukleasen DNA-Abschnitte entfernt (Deletion), bzw durch  $\Omega$ -Elemente substituiert (Insertion), wodurch das jeweilige Gen inaktiviert wurde. Die auf diese Weise mutierten Gene wurden in konjugativ übertragbare Vektoren umkloniert und anschließend in den Stamm Pseudomonas sp. HR199 eingeführt. Durch geeignete Selektion wurden Transkonjuganten erhalten, die das jeweils funktionsfähige wildtyp-Gen gegen das neu eingebrachte inaktivierte Gen ausgetauscht hatten. Die so erhaltenen Insertions- und Deletionsmutanten wiesen nur noch das jeweils inaktivierte Gen auf. Auf diese Weise wurden sowohl Mutanten mit nur einem defekten Gen als auch Mehrfachmutanten, in denen mehrere Gene auf diese Weise inaktiviert wurden, erhalten. Diese Mutanten wurden für die Biotransformation von

- a) Eugenol zu Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure;
  - b) Coniferylalkohol zu Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure;
  - c) Coniferylaldehyd zu Ferulasäure, Vanillin und/oder Vanillinsäure;
- d) Ferulasäure zu Vanillin und/oder Vanillinsäure und
  - e) Vanillin zu Vanillinsäure eingesetzt.

5

10

15

#### Material und Methoden

#### Wachstumsbedingungen der Bakterien.

Stämme von Escherichia coli wurden bei 37°C in Luria-Bertani (LB) oder M9-Mine-5 ralmedium (Sambrook, J., E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.) angezogen. Stämme von Pseudomonas sp. wurden bei 30°C in Nutrient Broth (NB, 0,8%, wt/vol) oder in Mineralmedium (MM) (Schlegel, H. G. et al. 1961. Arch. Mikrobiol. 38:209-222) bzw. HR-Mineralmedium (HR-MM) 10 (Rabenhorst, J. 1996. Appl. Microbiol. Biotechnol. 46:470-474.) angezogen. Ferulasäure, Vanillin, Vanillinsäure und Protocatechusäure wurden in Dimethylsulfoxid gelöst, und dem jeweiligen Medium in einer Endkonzentration von 0.1% (wt/vol) zugesetzt. Eugenol wurde dem Medium direkt in einer Endkonzentration von 0.1% (vol/vol) zugesetzt, bzw. in den Deckel von MM-Agarplatten auf Filterpapier 15 (Rundfilter 595, Schleicher & Schuell, Dassel, Deutschland) appliziert. Bei der Anzucht von Transkonjuganten und Mutanten von Pseudomonas sp. wurde Tetracyclin, Kanamycin, und Gentamycin in Endkonzentrationen von 25 µg/ml bzw. 100 μg/ml bzw. 7,5 μg/ml eingesetzt.

# Qualitativer und quantitativer Nachweis von Stoffwechselintermediaten in Kulturüberständen.

Kulturüberstände wurden direkt, bzw. nach Verdünnung mit H<sub>2</sub>O-bidest. mittels Hochdruck-Flüssigkeits-Chromatographie (Knauer-HPLC) analysiert. Die Chromatographie erfolgte an Nucleosil-100 C18 (7 μm, 250 x 4 mm). Als Lösungsmittel diente 0.1% (vol/vol) Ameisensäure und Acetonitril. Der verwendete Gradient zur Elution der Substanzen verlief wie folgt:

 $00:00 - 06:30 \rightarrow 26\%$  Acetonitril

 $06:30 - 08:00 \rightarrow 100\%$  Acetonitril

30  $08:00 - 12:00 \rightarrow 100\%$  Acetonitril

25

 $12:00 - 13:00 \rightarrow 26\%$  Acetonitril

13:00 - 18:00 → 26% Acetonitril

# Reinigung der Vanillin-Dehydrogenase-II.

Die Aufreinigung erfolgte bei 4°C.

#### 5 Rohextrakt.

10

15

20

Auf Eugenol angezogene Zellen von *Pseudomonas* sp. HR199 wurden in 10 mM Natriumphosphat-Puffer, pH 6.0 gewaschen, im gleichen Puffer resuspendiert und durch zweimalige Passage einer French-Presse (Amicon, Silver Spring, Maryland, USA) bei einem Druck von 1000 psi aufgeschlossen. Das Zellhomogenat wurde einer Ultrazentrifugation (1 h, 100 000 x g, 4°C) unterzogen, wodurch die lösliche Fraktion des Rohextraktes als Überstand erhalten wurde.

# Anionenaustauschchromatographie an DEAE-Sephacel.

Die lösliche Fraktion des Rohextraktes wurde über Nacht gegen 10 mM Natriumphosphat-Puffer, pH 6.0 dialysiert. Das Dialysat wurde auf eine mit 10 mM Natriumphosphat-Puffer, pH 6.0 äquilibrierte DEAE-Sephacel-Säule (2,6 cm x 35 cm, Bettvolumen[BV]: 186 ml) mit einer Durchflußrate von 0.8 ml/min aufgetragen. Die Säule wurde mit zwei BV 10 mM Natriumphosphat-Puffer, pH 6.0 gespült. Die Elution der Vanillin-Dehydrogenase-II (VDH-II) erfolgte mit einem linearen Salzgradient von 0 bis 400 mM NaCl in 10 mM Natriumphosphat-Puffer, pH 6.0 (750 ml). Es wurden 10 ml-Fraktionen aufgefangen. Fraktionen mit hoher VDH-II-Aktivität wurden zum DEAE-Pool vereinigt.

#### Bestimmung der Vanillin-Dehydrogenase-Aktivität.

Die Bestimmung der VDH-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test. Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.1 mmol Kalium-Phosphat (pH 7.1), 0.125 μmol Vanillin, 0.5 μmol NAD, 1.2 μmol Pyruvat (Na-Salz), Lactat-Dehydrogenase (1 U; aus Schweineherz ) und Enzymlösung. Die Oxidation von Vanillin wurde bei λ = 340 nm verfolgt (ε<sub>Vanillin</sub> = 11,6 cm<sup>2</sup>/μmol).

Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1 μmol Vanillin pro Minute umsetzt. Die Proteinkonzentrationen in

den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. 193:265-275) bestimmt.

# Bestimmung der Coniferylalkohol-Dehydrogenase-Aktivität.

Die Bestimmung der CADH-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test nach Jaeger et al. (Jaeger, E., L. Eggeling und H. Sahm. 1981. Current Microbiology. 6:333-336). Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.2 mmol Tris/HCl (pH 9.0), 0.4 μmol Coniferylalkohol, 2 μmol NAD, 0.1 mmol Semicarbazid und Enzymlösung. Die Reduktion von NAD wurde bei λ = 340 nm verfolgt (ε = 6.3 cm²/μmol). Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1 μmol Substrat pro Minute umsetzt. Die Proteinkonzentrationen in den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. 193:265-275) bestimmt.

15

20

30

# Bestimmung der Coniferylaldehyd-Dehydrogenase-Aktivität.

Die Bestimmung der CALDH-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test. Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.1 mmol Tris/HCl (pH 8.8), 0.08  $\mu$ mol Coniferylaldehyd, 2.7  $\mu$ mol NAD und Enzymlösung. Die Oxidation von Coniferylaldehyd zu Ferulasäure wurde bei  $\lambda$  = 400 nm verfolgt ( $\epsilon$  = 34 cm²/ $\mu$ mol). Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1  $\mu$ mol Substrat pro Minute umsetzt. Die Proteinkonzentrationen in den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. **193:**265-275) bestimmt

25 bestimmt.

# Bestimmung der Ferulasäure-CoA-Synthetase (Ferulasäure-Thiokinase)-Aktivität.

Die Bestimmung der FCS-Aktivität erfolgte bei 30°C durch einen optisch enzymatischen Test, modifiziert nach Zenk et al. (Zenk et al. 1980. Anal. Biochem. 101:182-187). Der Reaktionsansatz mit einem Volumen von 1 ml enthielt 0.09 mmol Kalium-

Phosphat (pH 7.0), 2.1  $\mu$ mol MgCl<sub>2</sub>, 0.7  $\mu$ mol Ferulasäure, 2  $\mu$ mol ATP, 0.4  $\mu$ mol Coenzym A und Enzymlösung. Die Entstehung des CoA-Esters aus Ferulasäure wurde bei  $\lambda = 345$  nm verfolgt ( $\epsilon = 10$  cm<sup>2</sup>/ $\mu$ mol). Die Enzymaktivität wurde in Einheiten (U) angegeben, wobei 1 U der Enzymmenge entspricht, die 1  $\mu$ mol Substrat pro Minute umsetzt. Die Proteinkonzentrationen in den Proben wurden nach Lowry et al. (Lowry, O. H., N. J. Rosebrough, A. L. Farr und R. J. Randall. 1951. J. Biol. Chem. 193:265-275) bestimmt.

#### Electrophoretische Methoden.

5

10

15

20

25

30

Die Auftrennung von proteinhaltigen Extrakten erfolgte in 7.4% (wt/vol) Polyacrylamidgelen unter nativen Bedingungen nach der Methode von Stegemann et al. (Stegemann et al. 1973. Z. Naturforsch. 28c:722-732) und unter denaturierenden Bedingungen in 11.5% (wt/vol) Polyacrylamidgelen nach der Methode von Laemmli (Laemmli, U. K. 1970. Nature (London) 227:680-685). Zur unspezifischen Proteinfärbung wurde Serva Blue R verwendet. Zur spezifischen Anfärbung der Coniferylalkohol-, Coniferylaldehyd- und Vanillin-Dehydrogenase wurden die Gele für 20 min in 100 mM KP-Puffer (pH 7.0) umgepuffert und anschließend bei 30°C im gleichen Puffer dem 0.08% (wt/vol) NAD, 0.04% (wt/vol) p-Nitroblau-Tetrazolium-chlorid, 0.003% (wt/vol) Phenazine-Methosulfat und 1 mM des jeweiligen Substrates zugesetzt worden war inkubiert, bis entsprechende Farbbanden sichtbar wurden.

Transfer von Proteinen aus Polyacrylamidgelen auf PVDF-Membranen. Proteine wurden aus SDS-Polyacrylamidgelen mit Hilfe eines Semidry-Fastblot Gerätes (B32/33, Biometra, Göttingen, Deutschland) nach Herstellerangaben auf PVDF-Membranen (Waters-Milipore, Bedford, Mass., USA) übertragen.

## Bestimmung von N-terminalen Aminosäuresequenzen.

Die Bestimmung von N-terminalen Aminosäuresequenzen erfolgte mit Hilfe eines Protein Peptide Sequenzers (Typ 477 A, Applied Biosystems, Foster City, USA) und eines PTH-Analysers nach Herstellerangaben.

# Isolierung und Manipulation von DNA.

Die Isolierung von genomischer DNA erfolgte nach der Methode von Marmur (Marmur, J. 1961. J. Mol. Biol. 3:208-218). Die Isolierung und Analyse von anderer Plasmid-DNA bzw. von DNA-Restriktionsfragmenten erfolgte nach Standardmethoden (Sambrook, J. E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Habor, New York.).

#### Transfer von DNA.

5

Die Präparation und Transformation von kompetenten Escherichia coli-Zellen 10 erfolgte nach der Methode von Hanahan (Hanahan, D. 1983. J. Mol. Biol. 166:557-580). Konjugativer Plasmidtransfer zwischen Plasmid-tragenden Escherichia coli S17-1-Stämmen (Donor) und Pseudomonas sp.-Stämmen (Rezipient) erfolgte auf NB-Agarplatten nach der Methode von Friedrich et al. (Friedrich, B. et al. 1981. J. Bacteriol. 147:198-205), oder durch eine "Minikomplementations-Methode" auf 15 MM-Agarplatten mit 0.5% (wt/vol) Gluconat als C-Quelle und 25 µg/ml Tetracyclin oder 100 µg/ml Kanamycin. Dabei wurden Zellen des Rezipienten in einer Richtung als Impfstrich aufgetragen. Nach 5 min wurden dann Zellen der Donor-Stämme als Impfstriche aufgetragen, wobei der Rezipienten-Impfstrich gekreuzt wurde. Nach einer Inkubation für 48 h bei 30°C wuchsen die Transkonjuganten direkt hinter der 20 Kreuzungsstelle, wohingegen weder Donor- noch Rezipienten-Stamm zum Wachstum in der Lage war.

# Hybridisierungsexperimente.

DNA-Restriktionsfragmente wurden in einem 0.8% (wt/vol) Agarose-Gel in 50 mM

Tris- 50 mM Borsäure- 1.25 mM EDTA-Puffer (pH 8.5) elektrophoretisch aufgetrennt (Sambrook, J. E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Habor, New York.). Die Übertragung der denaturierten DNA aus dem Gel auf eine positiv geladene Nylonmembran (Porengröße: 0.45 μm, Pall Filtrationstechnik, Dreieich, Deutschland), die anschließende Hybridisierung mit biotinylierten, bzw.

Digoxigenin-markierten DNA-Sonden und die Herstellung dieser DNA-Sonden erfolgte nach Standardmethoden (Sambrook, J. E. F. Fritsch und T. Maniatis. 1989. Molecular cloning: a laboratory manual. 2. Aufl., Cold Spring Harbor Laboratory Press, Cold Spring Habor, New York.).

5

10

#### **DNA-Sequenzierung.**

Die Bestimmung von Nukleotidsequenzen erfolgte nach der Didesoxy-Kettenabbruch-Methode von Sanger et al. (Sanger et al. 1977. Proc. Natl. Acad. Sci. USA 74:5463-5467) "nicht-radioaktiv" mit einem "LI-COR DNA-Sequencer Modell 4000L" (LI-COR Inc., Biotechnology Division, Lincoln, NE, USA) unter Verwendung eines "Thermo Sequenase fluorescent labelled primer cycle sequencing kit with 7-deaza-dGTP" (Amersham Life Science, Amersham International pls, Little Chalfont, Buckinghamshire, England) jeweils nach Vorschrift des Herstellers.

Mit Hilfe von synthetischen Oligonukleotiden wurde nach der "Primer-hopping Strategie" von Strauss et al. (Strauss, E. C. et al. 1986. Anal. Biochem. **154:**353-360) sequenziert.

## Chemikalien, Biochemikalien und Enzyme.

Restriktionsenzyme, T4 DNA-Ligase, Lambda-DNA und Enzyme bzw. Substrate für die optisch enzymatischen Tests wurden von C. F. Boehringer & Söhne (Mannheim, Deutschland) oder von GIBCO/BRL (Eggenstein, Deutschland) bezogen. [γ-32P]ATP kam von Amersham/Buchler (Braunschweig, Deutschland). Oligonukleotide wurden von der Firma MWG-Biotech GmbH (Ebersberg, Deutschland) bezogen. Agarose vom Typ NA wurde von Pharmacia-LKB (Uppsala, Schweden) bezogen. Alle anderen Chemikalien waren von Haarmann & Reimer (Holzminden, Deutschland), E. Merck AG (Darmstadt, Deutschland), Fluka Chemie (Buchs,, Schweiz), Serva Feinbiochemica (Heidelberg, Deutschland) oder Sigma Chemie (Deisenhofen, Deutschland).

#### Beispiele

# Beispiel 1

10

15

5 Konstruktion von Omega-Elementen, die Resistenzen gegenüber Kanamycin (Ω Km) bzw. Gentamycin( ΩGm) vermitteln.

Für die Konstruktion des ΩKm-Elements wurde das 2099 bp *Bgl*I-Fragment des Transposons Tn5 (Auerswald E. A., G. Ludwig und H. Schaller. 1981. Cold Spring Harb. Symp. Quant. Biol. **45**:107-113; Beck E., G. Ludwig, E. A. Auerswald, B. Reiss und H. Schaller. 1982. Gene **19**:327-336; Mazodier P., P. Cossart, E. Giraud und F. Gasser. 1985. Nucleic Acids Res. **13**:195-205.) präparativ isoliert. Das Fragment wurde durch Behandlung mit der Nuklease Bal-31 auf ca. 990 bp verkürzt. Dieses Fragment, das nur noch das Kanamycin-Resistenzgen (codierend für eine Aminoglycosid-3'-O-Phosphotransferase) umfaßte, wurde anschließend mit *Sma*I geschnittener pSKsym-DNA (pBluescript SK<sup>-</sup>-Derivat, welches eine symetrisch aufgebaute multiple Klonierungsstelle [*Sal*I, *Hin*dIII, *Eco*RI, *Sma*I, *Eco*RI, *Hin*dIII, *Sal*I] enthält) ligiert. Aus dem resultierenden Plasmid konnte das ΩKm-Element als *Sma*I-, *Eco*RI-, *Hin*dIII- oder *Sal*I-Fragment reisoliert werden.

Plasmids pBBR1MCS-5 (Kovach M.E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop und K. M. Peterson. 1995. Gene 166:175-176.) präparativ isoliert und anschließend mit Mung Bean Nuklease (Abdauen von einzelsträngigen DNA-Molekülenden) behandelt. Dieses Fragment, das nur noch das Gentamycin-Resistenzgen (codierend für eine Gentamycin-3-Acetyltransferase) umfaßte, wurde anschließend mit *Sma*I geschnittener pSKsym-DNA (s.o.) ligiert. Aus dem resultierenden Plasmid konnte das ΩGm-Element als *Sma*I-, *Eco*RJ-, *Hin*dIII- oder *Sal*I-Fragment reisoliert werden.

# Beispiel 2

Klonierung der Gene aus *Pseudomonas* sp. HR199 (DSM7063), die durch Insertion von  $\Omega$ -Elementen oder durch Deletionen inaktiviert werden sollten.

- Die separaten Klonierungen der Gene fcs, ech, vdh und aat erfolgte ausgehend von den E. coli S17-1 Stämmen DSM 10439 und DSM 10440 mit den Plasmiden pE207 und pE5-1 (siehe EP-A 0845532). Aus diesen Plasmiden wurden die angegebenen Fragmente präparativ isoliert und wie im weiteren beschrieben behandelt:
- Für die Klonierung des fcs-Gens wurde das 2350 bp große Sall/EcoRI-Fragment des Plasmids pE207 und das 3700 bp große EcoRI/Sall-Fragment des Plasmids pE5-1 zusammen in pBluescript SK in einer Weise kloniert, daß beide Fragmente über die EcoRI-Enden miteinander verbunden waren. Ausgehend von dem resultierenden Hybridplasmid wurde das 6050 bp Sall-Fragment präparativ isoliert und durch Behandlung mit der Nuklease Bal-31 auf ca. 2480 bp verkürzt. Anschließend wurden an die Fragment-Enden Pstl-Linker ligiert und das Fragment nach Pstl-Verdauung in pBluescript SK kloniert (pSKfcs). Nach Transformation von E. coli XL1-Blue wurden Klone erhalten, die das fcs-Gen exprimierten und eine FCS-Aktivität von 0.2 U/mg Protein aufwiesen.

20

25

30

Für die Klonierung des *ech*-Gens wurde das 3800 bp große *Hin*dIII/*Eco*RI-Fragment des Plasmids pE207 präparativ isoliert und durch Behandlung mit der Nuklease Bal-31 auf ca. 1470 bp verkürzt. Anschließend wurden an die Fragment-Enden *Eco*RI-Linker ligiert und das Fragment nach *Eco*RI-Verdauung in pBluescript SK<sup>-</sup> kloniert (pSK*ech*).

Für die Klonierung des *vdh*-Gens wurde das 2350 bp große *Sall/Eco*RI-Fragment des Plasmids pE207 präparativ isoliert. Nach Klonierung in pBluescript SK<sup>-</sup> wurde das Fragment mit Hilfe eines Exonuklease III/ Mung Bean Nukleasesystems einseitig um ca. 1530 bp verkürzt. Anschließend wurde an das Fragmentende ein *Eco*RI-

Linker ligiert und das Fragment nach *Eco*RI-Verdauung in pBluescript SK<sup>-</sup> kloniert (pSK*vdh*). Nach Transformation von *E. coli* XL1-Blue wurden Klone erhalten, die das *vdh*-Gen exprimierten und eine VDH-Aktivität von 0.01 U/mg Protein aufwiesen.

5

10

15

20

25

30

Für die Klonierung des *aat*-Gens wurde das 3700 bp große *EcoRI/Sal*I-Fragment des Plasmids pE5-1 präparativ isoliert und durch Behandlung mit der Nuklease Bal-31 auf ca. 1590 bp verkürzt. Anschließend wurden an die Fragment-Enden *EcoRI*-Linker ligiert und das Fragment nach *EcoRI*-Verdauung in pBluescript SK<sup>-</sup> kloniert (pSK*aat*).

# Beispiel 3

Inaktivierung der oben beschriebenen Gene durch Insertion von  $\Omega$ -Elementen, bzw durch Deletion von Teilbereichen dieser Gene.

Das Plasmid pSKfcs, welches das fcs-Gen enthielt wurde mit BssHII verdaut, wodurch ein 1290 bp großes Fragment aus dem fcs-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des fcs-Gens ( $fcs\Delta$ ) (siehe Abb. 1i und 2i) kloniert in pBluescript SK $^-$  (pSK $fcs\Delta$ ) erhalten. Darüber hinaus wurden nach Herausschneiden des Fragments die Omega-Elemente  $\Omega$ Km und  $\Omega$ Gm an dessen Stelle einligiert. Dadurch entstanden die  $\Omega$ -inaktivierten Derivate des fcs-Gens ( $fcs\Omega$ Km, siehe Abb. 1g und 2g) und ( $fcs\Omega$ Gm, siehe Abb. 1h und 2h) kloniert in pBluescript SK $^-$  (pSK $fcs\Omega$ Km und pSK $fcs\Omega$ Gm). In Rohextrakten der erhaltenen E. coli Klone, deren Hybridplasmide ein durch Deletion bzw.  $\Omega$ -Element-Insertion inaktiviertes fcs-Gen aufwiesen, konnte keine FCS-Aktivität nachgewiesen werden.

Das Plasmid pSKech, welches das ech-Gen enthielt, wurde mit Nrul verdaut, wodurch ein 53 bp und ein 430 bp großes Fragment aus dem ech-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des ech-Gens (echΔ, siehe Abb. 11 und 21) kloniert in pBluescript SK<sup>-</sup> (pSKechΔ) erhalten. Darüber

hinaus wurden nach Herausschneiden der Fragmente die Omega-Elemente  $\Omega$ Km und  $\Omega$ Gm an deren Stelle einligiert. Dadurch entstanden die  $\Omega$ -inaktivierten Derivate des ech-Gens ( $ech\Omega$ Km und  $ech\Omega$ Gm) kloniert in pBluescript SK (pSK $ech\Omega$ Km und pSK $ech\Omega$ Gm).

5

10

15

20

25

Das Plasmid pSKvdh, welches das vdh-Gen enthielt wurde mit BssHII verdaut, wodurch ein 210 bp großes Fragment aus dem vdh-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des vdh-Gens ( $vdh\Delta$ , siehe Abb. 10 und 20) kloniert in pBluescript SK $^-$  (pSK $vdh\Delta$ ) erhalten. Darüber hinaus wurden nach Herausschneiden des Fragments die Omega-Elemente  $\Omega$ Km und  $\Omega$ Gm an dessen Stelle einligiert. Dadurch entstanden die  $\Omega$ -inaktivierten Derivate des vdh-Gens ( $vdh\Omega$ Km und  $vdh\Omega$ Gm) kloniert in pBluescript SK $^-$  (pSK $vdh\Omega$ Km, siehe Abb. 1m und 2m) und (pSK $vdh\Omega$ Gm, siehe Abb. 1n und 2n). In Rohextrakten der erhaltenen E. coli Klone, deren Hybridplasmide ein durch Deletion bzw.  $\Omega$ -Element-Insertion inaktiviertes vdh-Gen aufwiesen, konnte keine VDH-Aktivität nachgewiesen werden.

Das Plasmid pSKaat, welches das aat-Gen enthielt wurde mit BssHII verdaut, wodurch ein 59 bp großes Fragment aus dem aat-Gen herausgeschnitten wurde. Nach Religation wurde das Deletions-Derivat des aat-Gens (aat $\Delta$ , siehe Abb. 1r und 2r) kloniert in pBluescript SK (pSKaat $\Delta$ ) erhalten. Darüber hinaus wurden nach Herausschneiden des Fragments die Omega-Elemente  $\Omega$ Km und  $\Omega$ Gm an dessen Stelle einligiert. Dadurch entstanden die  $\Omega$ -inaktivierten Derivate des aat-Gens (aat  $\Omega$ Km, siehe Abb. 1p und 2p) und (aat $\Omega$ Gm, siehe Abb. 1q und 2q) kloniert in pBluescript SK (pSKaat $\Omega$ Km und pSKaat $\Omega$ Gm).

# Beispiel 4

5

10

15

20

Umklonieren der durch  $\Omega$ -Elemente inaktivierten Gene in das konjugativ übertragbare "Suizid-Plasmid" pSUP202.

Um die durch Ω-Elemente inaktivierten Gene in *Pseudomonas* sp. HR199 gegen die intakten Gene austauschen zu können, benötigt man einen Vektor, der einerseits in Pseudomonaden übertragen werden kann (konjugativ übertragbare Plasmide), andererseits dort jedoch nicht repliziert werden kann und somit in Pseudomonaden instabil ist ("Suizid-Plasmid"). DNA-Abschnitte, die mit Hilfe eines solchen Plasmidsystems in Pseudomonaden übertragen werden, können nur erhalten bleiben, wenn sie durch homologe Rekombination (RecA-abhängige Rekombination) in das Genom der Bakterienzelle integriert werden. Im vorliegenden Fall wurde das "Suizid-Plasmid" pSUP202 (Simon et al. 1983. *In*: A. Pühler. Molecular genetics of the bacteria-plant interaction. Springer Verlag, Berlin, Heidelberg, New York, S. 98-106.) eingesetzt.

Die inaktivierten Gene  $fcs\Omega$ Km und  $fcs\Omega$ Gm wurden nach PstI-Verdauung aus den Plasmiden pSK $fcs\Omega$ Km und pSK $fcs\Omega$ Gm isoliert und mit PstI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren Hybridplasmid (pSUP $fcs\Omega$ Km) das inaktivierte Gen  $fcs\Omega$ Km enthielt. Das entsprechende Hybridplasmid (pSUP $fcs\Omega$ Gm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen  $fcs\Omega$ Gm.

25

30

Die inaktivierten Gene  $ech\Omega$ Km und  $ech\Omega$ Gm wurden nach EcoRI-Verdauung aus den Plasmiden pSK $ech\Omega$ Km und pSK $ech\Omega$ Gm isoliert und mit EcoRI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren

Hybridplasmid (pSUP $ech\Omega$ Km) das inaktivierte Gen  $ech\Omega$ Km enthielt. Das entsprechende Hybridplasmid (pSUP $ech\Omega$ Gm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen  $ech\Omega$ Gm.

Die inaktivierten Gene  $vdh\Omega$ Km und  $vdh\Omega$ Gm wurden nach EcoRI-Verdauung aus den Plasmiden pSK $vdh\Omega$ Km und pSK $vdh\Omega$ Gm isoliert und mit EcoRI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren Hybridplasmid (pSUP $vdh\Omega$ Km) das inaktivierte Gen  $vdh\Omega$ Km enthielt. Das entsprechende Hybridplasmid (pSUP $vdh\Omega$ Gm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen  $vdh\Omega$ Gm.

Die inaktivierten Gene  $aat\Omega$ Km und  $aat\Omega$ Gm wurden nach EcoRI-Verdauung aus den Plasmiden pSK $aat\Omega$ Km und pSK $aat\Omega$ Gm isoliert und mit EcoRI geschnittener pSUP202 DNA ligiert. Die Ligationsansätze wurden nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium mit Kanamycin bzw. Gentamycin. Es wurden Kanamycin-resistente Transformanden erhalten, deren Hybridplasmid (pSUP $aat\Omega$ Km) das inaktivierte Gen  $aat\Omega$ Km enthielt. Das entsprechende Hybridplasmid (pSUP $aat\Omega$ Gm) der Gentamycin- resistenten Transformanden enthielt das inaktivierte Gen  $aat\Omega$ Gm.

## Beispiel 5

5

10

15

20

30

Umklonieren der durch Deletion inaktivierten Gene in das konjugativ übertragbare "Suizid-Plasmid" mit "sacB-Selektionssystem" pHE55.

Um die durch Deletion inaktivierten Gene in *Pseudomonas* sp. HR199 gegen die intakten Gene austauschen zu können, benötigt man einen Vektor, der die schon für pSUP202 beschriebenen Eigenschaften aufweist. Da im Gegensatz zu den durch  $\Omega$ -Element inaktivierten Genen bei durch Deletion inaktivierten Genen keine Selek-

5

10

15

20

25

30

tionsmöglichkeit (keine Antibiotika-Resistenz) für den erfolgten Austausch der Gene in Pseudomonas sp. HR199 besteht, mußte ein anderes Selektionssystem zur Anwendung kommen. Bei dem "sacB-Selektionssystem" wird das auszutauschende, durch Deletion inaktivierte Gen in einem Plasmid kloniert, welches neben einem Antibiotika-Resistenzgen auch über das sacB-Gen verfügt. Nach konjugativer Übertragung dieses Hybridplasmids in einen Pseudomonaden wird das Plasmid durch homologe Rekombination an der Stelle in das Genom integriert, an der sich das intakte Gen befindet (erster "Cross over"). Auf diese Weise entsteht ein "heterogenoter" Stamm, der sowohl über ein intaktes als auch über ein durch Deletion inaktiviertes Gen verfügt, welche durch die pHE55-DNA voneinander getrennt sind. Diese Stämme weisen die durch den Vektor codierte Resistenz auf und besitzen darüber hinaus ein aktives sacB-Gen. Durch ein zweites homologes Rekombinationsereignis (zweiter "Cross over"), soll nun die pHE55-DNA zusammen mit dem intakten Gen aus der genomischen DNA ausgegliedert werden. Durch dieses Rekombinationsereignis entsteht ein Stamm, der nur noch über das inaktivierte Gen verfügt. Darüber hinaus kommt es zum Verlust der pHE55-codierten Antibiotika-Resistenz und des sacB-Gens. Streicht man Stämme auf Saccharose-haltigen Medien aus, werden Stämme die das sacB-Gen exprimieren im Wachstum gehemmt, da das Genprodukt Saccharose zu einem Polymer umsetzt, welches im Periplasma der Zellen akkumuliert wird. Zellen, die durch das zweite Rekombinationsereignis das sacB-Gen nicht mehr tragen, werden somit nicht im Wachstum gehemmt. Um eine phänotypische Selektionsmöglichkeit auf die Integration des durch Deletion inaktivierten Gens zu besitzen, tauscht man dieses nicht gegen ein intaktes Gen aus, sondern man bedient sich eines Stammes, in dem das auszutauschende Gen bereits durch Insertion eines  $\Omega$ -Elements "markiert" vorliegt. Bei erfolgreichem Austausch verliert der resultierende Stamm die durch das Ω-Element codierte Antibiotika-Resistenz.

Das inaktivierte Gen fcsΔ wurden nach PstI-Verdauung aus dem Plasmid pSKfcsΔ isoliert und mit PstI geschnittener pHE55 DNA ligiert. Der Ligationsansatz wurde nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-

Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE $fcs\Delta$ ) das inaktivierte Gen  $fcs\Delta$  enthielt.

Das inaktivierte Gen  $ech\Delta$  wurden nach EcoRI-Verdauung aus dem Plasmid pSK $ech\Delta$  isoliert und mit Mung Bean Nuklease behandelt (Erzeugung von glatten Enden ["blunt ends"]). Das Fragment wurde mit BamHI geschnittener und Mung Bean Nuklease behandelter pHE55 DNA ligiert. Der Ligationsansatz wurde nach Ecoli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE $ech\Delta$ ) das inaktivierte Gen  $ech\Delta$  enthielt.

Das inaktivierte Gen  $vdh\Delta$  wurden nach EcoRI-Verdauung aus dem Plasmid pSK $vdh\Delta$  isoliert und mit Mung Bean Nuklease behandelt. Das Fragment wurde mit BamHI geschnittener und Mung Bean Nuklease behandelter pHE55 DNA ligiert. Der Ligationsansatz wurde nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE $vdh\Delta$ ) das inaktivierte Gen  $vdh\Delta$  enthielt.

Das inaktivierte Gen  $aat\Delta$  wurden nach EcoRI-Verdauung aus dem Plasmid pSK $aat\Delta$  isoliert und mit Mung Bean Nuklease behandelt. Das Fragment wurde mit BamHI geschnittener und Mung Bean Nuklease behandelter pHE55 DNA ligiert. Der Ligationsansatz wurde nach E. coli S17-1 transformiert. Die Selektion erfolgte auf Tetracyclin-haltigem LB-Medium. Es wurden Tetracyclin-resistente Transformanden erhalten, deren Hybridplasmid (pHE $aat\Delta$ )das inaktivierte Gen  $aat\Delta$  enthielt.

5

10

15

20

Erzeugung von Mutanten des Stammes *Pseudomonas* sp. HR199, bei denen spezifisch Gene des Eugenol-Katabolismuses durch Insertion eines  $\Omega$ -Elementes inaktiviert wurden.

Der Stamm Pseudomonas sp. HR199 wurde als Rezipient in Konjugationsexperimenten eingesetzt, bei denen Stämme von E. coli S17-1 als Donoren eingesetzt wurden, die die unten aufgeführten Hybridplasmide von pSUP202 enthielten. Die Transkonjuganten wurden auf Gluconat-haltigem Mineralmedium selektiert, welches das dem  $\Omega$ -Element entsprechende Antibiotikum enthielt. "Homogenote" (Austausch des intakten Gens gegen das durch  $\Omega$ -Element-Insertion inaktivierte Gen durch doppeltes "Cross over") und "heterogenote" (Integration des Hybridplasmids in das Genom durch einfachen "Cross over") Transkonjuganten konnten anhand der durch pSUP202 codierten Tetracyclin-Resistenz unterschieden werden.

15

20

25

30

5

10

Die Mutanten Pseudomonas sp. HR199  $fcs\Omega$ Km und Pseudomonas sp. HR199  $fcs\Omega$  Gm wurden nach Konjugation von Pseudomonas sp. HR199 mit E. coli S17-1 (pSUP $fcs\Omega$ Km) bzw. E. coli S17-1 (pSUP $fcs\Omega$ Gm) erhalten. Der Austausch des intakten fcs-Gens gegen das durch  $\Omega$ Km bzw.  $\Omega$ Gm inaktivierte Gen ( $fcs\Omega$ Km bzw.  $fcs\Omega$ Gm) wurde mittels DNA-Sequenzierung verifiziert.

Die Mutanten Pseudomonas sp. HR199  $ech\Omega$ Km und Pseudomonas sp. HR199  $ech\Omega$  Gm wurden nach Konjugation von Pseudomonas sp. HR199 mit E. coli S17-1 (pSUP $ech\Omega$ Km) bzw. E. coli S17-1 (pSUP $ech\Omega$ Gm) erhalten. Der Austausch des intakten ech-Gens gegen das durch  $\Omega$ Km bzw.  $\Omega$ Gm inaktivierte Gen ( $ech\Omega$ Km bzw.  $ech\Omega$ Gm) wurde mittels DNA-Sequenzierung verifiziert.

Die Mutanten *Pseudomonas* sp. HR199  $vdh\Omega$ Km und *Pseudomonas* sp. HR199  $vdh\Omega$ Gm wurden nach Konjugation von *Pseudomonas* sp. HR199 mit *E. coli* S17-1 (pSUP $vdh\Omega$ Km) bzw. *E. coli* S17-1 (pSUP $vdh\Omega$ Gm) erhalten. Der Austausch des

intakten vdh-Gens gegen das durch  $\Omega$ Km bzw.  $\Omega$ Gm inaktivierte Gen ( $vdh\Omega$ Km bzw.  $vdh\Omega$ Gm) wurde mittels DNA-Sequenzierung verifiziert.

Die Mutanten *Pseudomonas* sp. HR199  $aat\Omega$ Km und *Pseudomonas* sp. HR199  $aat\Omega$  Gm wurden nach Konjugation von *Pseudomonas* sp. HR199 mit *E. coli* S17-1 (pSUP $aat\Omega$ Km) bzw. *E. coli* S17-1 (pSUP $aat\Omega$ Gm) erhalten. Der Austausch des intakten aat-Gens gegen das durch  $\Omega$ Km bzw.  $\Omega$ Gm inaktivierte Gen ( $aat\Omega$ Km bzw.  $aat\Omega$ Gm) wurde mittels DNA-Sequenzierung verifiziert.

5

- Die Mutante *Pseudomonas* sp. HR199 fcsΩKmvdhΩGm wurden nach Konjugation von *Pseudomonas* sp. HR199 fcsΩKm mit E. coli S17-1 (pSUPvdhΩGm) erhalten. Der Austausch des intakten vdh-Gens gegen das durch ΩGm inaktivierte Gen (vdhΩGm) wurde mittels DNA-Sequenzierung verifiziert.
- Die Mutante *Pseudomonas* sp. HR199 *vdh*ΩKm*aat*ΩGm wurden nach Konjugation von *Pseudomonas* sp. HR199 *vdh*ΩKm mit *E. coli* S17-1 (pSUP*aat*ΩGm) erhalten. Der Austausch des intakten *aat*-Gens gegen das durch ΩGm inaktivierte Gen (*aat*Ω Gm) wurde mittels DNA-Sequenzierung verifiziert.
- Die Mutante *Pseudomonas* sp. HR199 *vdh*ΩKm*ech*ΩGm wurden nach Konjugation von *Pseudomonas* sp. HR199 *vdh*ΩKm mit *E. coli* S17-1 (pSUP*ech*ΩGm) erhalten. Der Austausch des intakten *ech*-Gens gegen das durch ΩGm inaktivierte Gen (*ech*Ω Gm) wurde mittels DNA-Sequenzierung verifiziert.

# **Beispiel 7**

5

10

15

Erzeugung von Mutanten des Stammes *Pseudomonas* sp. HR199, bei denen spezifisch Gene des Eugenol-Katabolismuses durch Deletion eines Teilbereiches inaktiviert wurden.

Die Stämme Pseudomonas sp. HR199 fcs $\Omega$ Km, Pseudomonas sp. HR199 ech $\Omega$ Km, Pseudomonas sp. HR199 vdh $\Omega$ Km und Pseudomonas sp. HR199 aat $\Omega$ Km wurden als Rezipient in Konjugationsexperimenten eingesetzt, bei denen Stämme von E. coli S17-1 als Donoren eingesetzt wurden, die die unten aufgeführten Hybridplasmide von pHE55 enthielten. Die "heterogenoten" Transkonjuganten wurden auf Gluconathaltigem Mineralmedium selektiert, welches neben Tetracyclin (pHE55 codierte Resistenz) das dem  $\Omega$ -Element entsprechende Antibiotikum enthielt. Nach Ausstreichen auf Saccharose-haltigem Mineralmedium wurden Transkonjuganten erhalten, die durch ein zweites Rekombinationsereignis (zweiter "Cross over") die Vektor-DNA eliminiert hatten. Durch Ausstreichen auf Mineralmedium ohne Antibiotika bzw. mit dem  $\Omega$ -Element entsprechenden Antibiotikum konnten die Mutanten erkannt werden, bei denen das durch  $\Omega$ -Element inaktivierte Gen gegen das durch Deletion inaktivierte Gen ausgetauscht worden war (keine Antibiotika-Resistenz).

- Die Mutante *Pseudomonas* sp. HR199 fcsΔ wurde nach Konjugation von *Pseudomonas* sp. HR199 fcsΩKm mit E. coli S17-1 (pHEfcsΔ) erhalten. Der Austausch des durch ΩKm inaktivierten Gens (fcsΩKm) gegen das durch Deletion inaktivierte Gen (fcsΔ) wurde mittels DNA-Sequenzierung verifiziert.
- Die Mutanten *Pseudomonas* sp. HR199 echΔ wurde nach Konjugation von *Pseudomonas* sp. HR199 echΩKm mit *E. coli* S17-1 (pHEechΔ) erhalten. Der Austausch des durch ΩKm inaktivierten Gens (echΩKm) gegen das durch Deletion inaktivierte Gen (echΔ) wurde mittels DNA-Sequenzierung verifiziert.

WO 00/26355 PCT/EP99/07952

Die Mutanten *Pseudomonas* sp. HR199  $vdh\Delta$  wurde nach Konjugation von *Pseudomonas* sp. HR199  $vdh\Omega$ Km mit *E. coli* S17-1 (pHE $vdh\Delta$ ) erhalten. Der Austausch des durch  $\Omega$ Km inaktivierten Gens ( $vdh\Omega$ Km) gegen das durch Deletion inaktivierte Gen ( $vdh\Delta$ ) wurde mittels DNA-Sequenzierung verifiziert.

5

Die Mutanten *Pseudomonas* sp. HR199  $aat\Delta$  wurde nach Konjugation von *Pseudomonas* sp. HR199  $aat\Omega$ Km mit *E. coli* S17-1 (pHE $aat\Delta$ ) erhalten. Der Austausch des durch  $\Omega$ Km inaktivierten Gens ( $aat\Omega$ Km) gegen das durch Deletion inaktivierte Gen ( $aat\Delta$ ) wurde mittels DNA-Sequenzierung verifiziert.

10

#### Beispiel 8

Biotransformation von Eugenol zu Vanillin mit der Mutante *Pseudomonas* sp. HR199 vdhΩKm.

Der Stamm *Pseudomonas* sp. HR199 *vdh*ΩKm wurde in 50 ml HR-MM mit 6 mM Eugenol bis zu einer optischen Dichte von ca. OD600nm = 0.6 angezogen. Nach 17 h waren 2.9 mM Vanillin, 1.4 mM Ferulasäure und 0.4 mM Vanillinsäure im Kultur-überstand nachweisbar.

# 20 Beispiel 9

25

Biotransformation von Eugenol zu Ferulasäure mit der Mutante *Pseudomonas* sp. HR199  $vdh\Omega Gmaat\Omega Km$ .

Der Stamm *Pseudomonas* sp. HR199 *vdh*ΩGmaatΩKm wurde in 50 ml HR-MM mit 6 mM Eugenol bis zu einer optischen Dichte von ca. OD600nm = 0.6 angezogen. Nach 18 h waren 1.9 mM Vanillin, 2.4 mM Ferulasäure und 0.6 mM Vanillinsäure im Kulturüberstand nachweisbar.

# Beispiel 10

Biotransformation von Eugenol zu Coniferylalkohol mit der Mutante Pseudomonas sp. HR199  $vdh\Omega Gmaat\Omega Km$ .

Der Stamm *Pseudomonas* sp. HR199 *vdh*ΩGmaatΩKm wurde in 50 ml HR-MM mit 6 mM Eugenol bis zu einer optischen Dichte von ca. OD600nm = 0.4 angezogen. Nach 15 h waren 1.7 mM Coniferylalkohol, 1.4 mM Vanillin, 1.4 mM Ferulasäure und 0.2 mM Vanillinsäure im Kulturüberstand nachweisbar.

# 10 Beispiel 11

Fermentative Produktion von natürlichem Vanillin aus Eugenol im 101 Fermenter mit der Mutante Pseudomonas sp. HR 199  $vdh\Omega$ Km.

Mit 100 ml einer 24 Stunden alten Vorkultur, die auf einer Schüttelmaschine (120 Upm) bei 32°C in einem auf pH 7,0 eingestellten Medium aus 12,5 g/l Glyzerin, 10 g/l Hefeextrakt und 0,37 g/l Essigsäure angezogen wurde, wurde der Produktionsfermenter beimpft. Der Fermenter enthielt 9,9 l Medium mit folgender Zusammensetzung: 1,5 g/l Hefeextrakt, 1,6 g/l KH<sub>2</sub>PO<sub>4</sub>, 0,2 g/l NaCl, 0,2 g/l MgSO<sub>4</sub>. Der pH-Wert wurde mit Natronlauge auf pH 7,0 eingestellt. Nach der Sterilisation wurde dem Medium 4 g Eugenol zugefügt. Die Temperatur betrug 32°C, die Belüftung 3 Nl/min und die Rührerdrehzahl 600 Upm. Der pH-Wert wurde mit Natronlauge bei pH 6,5 gehalten.

Vier Stunden nach dem Animpfen wurde mit der kontinuierlichen Zugabe von Eugenol begonnen, so daß am Ende der Fermentation nach 65 Stunden 255 g Eugenol zur Kultur gegeben worden waren. Außerdem wurden während der Fermentation 40 g Hefeextrakt zugefüttert. Die Konzentration an Eugenol lag am Ende der Fermentation bei 0,2 g/l. Der Gehalt an Vanillin betrug 2,6 g/l. Zusätzlich lagen noch 3,4 g/l Ferulasäure vor.

15

20

25

WO 00/26355 PCT/EP99/07952

Das so erhaltene Vanillin kann durch bekannte physikalische Verfahren wie Chromatographie, Destillation und/oder Extraktion isoliert und zur Herstellung natürlicher Aromen verwendet werden.

# 5 Erläuterungen zu den Figuren:

FIG. la bis 1r:

Gen- Strukturen zur Gewinnung von Organismen und Mutanten

aat\*: Teil des inaktivierten Gens der beta-Ketothiolase

10

calA\*: Teil des inaktivierten Gens der Coniferylalkohol-Dehydrogenase calB\*: Teil des inaktivierten Gens der Coniferylaldehyd-Dehydrogenase fcs\*: Teil des inaktivierten Gens der Ferulasäure-CoA Synthetase ech\*: Teil des inaktivierten Gens der Enoyl-CoA Hydratase-Aldolase vdh\*: Teil des inaktivierten Gens der Vanillin-Dehydrogenase

Die mit "\*" versehenen Restriktionsenzym-Schnittstellen kamen für die Konstruktion zum Einsatz, sind jedoch in dem resultierenden Konstrukt nicht mehr funktionsfähig.

20

15

	FIG. 2a: Nukleotidsequenz der Gen-Struktur calAΩKm
	FIG. 2b: Nukleotidsequenz der Gen-Struktur calAΩGm
	FIG. 2c: Nukleotidsequenz der Gen-Struktur calAA
	FIG. 2d: Nukleotidsequenz der Gen-Struktur calBΩKm
5	FIG. 2e: Nukleotidsequenz der Gen-Struktur calBΩGm
	FIG. 2f: Nukleotidsequenz der Gen-Struktur calB∆
	FIG. 2g: Nukleotidsequenz der Gen-Struktur fcsΩKm
	FIG. 2h: Nukleotidsequenz der Gen-Struktur fcsΩGm
	FIG. 2i: Nukleotidsequenz der Gen-Struktur fcs∆
10	FIG. 2j: Nukleotidsequenz der Gen-Struktur echΩKm
	FIG. 2k: Nukleotidsequenz der Gen-Struktur echΩGm
	FIG. 21: Nukleotidsequenz der Gen-Struktur ech∆
	FIG. 2m: Nukleotidsequenz der Gen-Struktur vdhΩKm
	FIG. 2n: Nukleotidsequenz der Gen-Struktur vdhΩGm
15	FIG. 20: Nukleotidsequenz der Gen-Struktur vdhA
	FIG. 2p: Nukleotidsequenz der Gen-Struktur aatΩKm
	FIG. 2a: Nukleotidsequenz der Gen-Struktur aatΩGm

FIG. 2r: Nukleotidsequenz der Gen-Struktur aat \( \Delta \)

5

15

# Patentansprüche

- 1. Transformierter und/oder mutagenisierter ein- oder mehrzelliger Organismus, der dadurch gekennzeichnet ist, daß Enzyme des Eugenol- und/oder Ferulasäure-Katabolismus derart inaktiviert sind, daß eine Akkumulation der Intermediate Coniferylalkohol, Coniferylaldehyd, Ferulasäure, Vanillin und/oder Vanillinsäure erfolgt.
- Organismus nach Anspruch 1, dadurch gekennzeichnet, daß der Eugenolund/oder Ferulasäure-Katabolismus durch Ω-Element-Insertion oder Einführen von Deletionen in entsprechende Gene verändert ist.
  - 3. Organismus nach einem der Ansprüche 1 oder 2, dadurch gekennzeichnet, daß ein oder mehrere Gene, die für die Enzyme Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA Synthetasen, Enoyl-CoA Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen Enzyme kodieren, verändert und/oder inaktiviert sind.
- 4. Organismus nach einem der Ansprüche 1 bis 3, dadurch gekennzeichnet, daß er einzellig, vorzugsweise ein Mikroorganismus oder eine pflanzliche oder eine tierische Zelle ist.
- 5. Organismus nach einem der Ansprüche 1 bis 4, dadurch gekennzeichnet, daß er ein Bakterium, vorzugsweise eine *Pseudomonas*-Art ist.
- Gen-Strukturen, bei denen die für die Enzyme Coniferylalkohol-Dehydrogenasen, Coniferylaldehyd-Dehydrogenasen, Ferulasäure-CoA Synthetasen, Enoyl-CoA Hydratasen-Aldolasen, beta-Ketothiolasen, Vanillin-Dehydrogenasen oder Vanillinsäure-Demethylasen oder zweier oder mehrerer dieser Enzyme kodierenden Nukleotidsequenzen verändert und/oder inaktiviert sind.

- 7. Gen-Strukturen mit den in Figur 1a bis 1r angegebenen Strukturen.
- 8. Gen-Strukturen mit den in Figur 2a bis 2r angegebenen Sequenzen.

5

- 9. Vektoren enthaltend wenigstens eine Gen-Struktur nach einem der Ansprüche6 bis 8.
- 10. Transformierter Organismus nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß er wenigstens einen Vektor gemäß Anspruch 9 enthält.
  - 11. Organismus nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß er wenigstens eine Gen-Struktur nach einem der Ansprüche 6 bis 8 an Stelle des jeweiligen intakten Gens im Genom integriert enthält.

15

12. Verfahren zur biotechnischen Herstellung von organischen Verbindungen, insbesondere von Alkoholen, Aldehyden und organischen Säuren, dadurch gekennzeichnet, daß ein Organismus nach einem der Ansprüche 1 bis 5 oder 10 bis 11 eingesetzt wird.

20

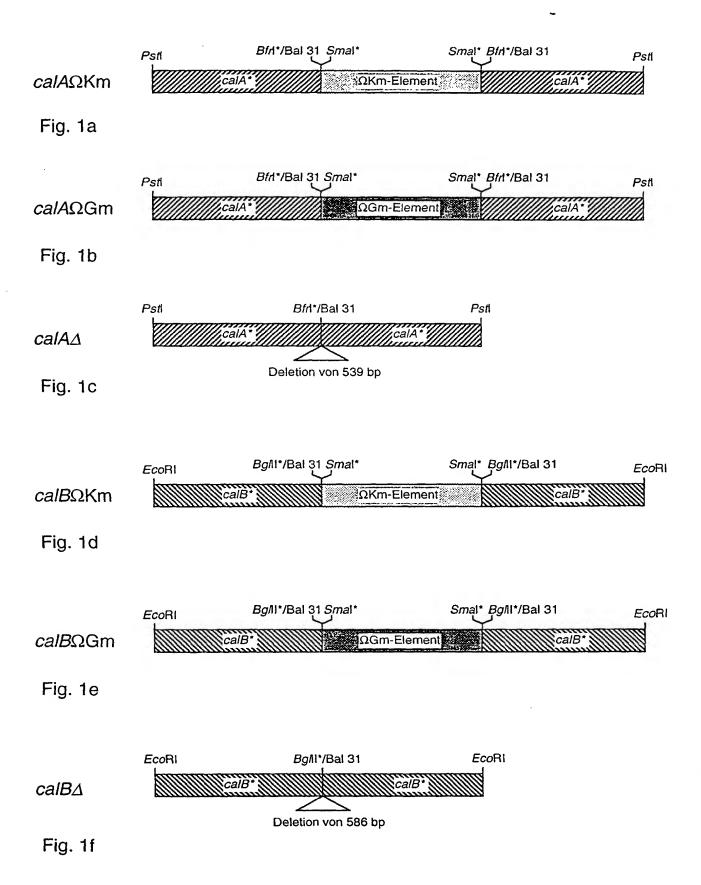
13. Verfahren zur Herstellung der Organismen nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, daß die Veränderung des Eugenol- und/oder Ferulasäure-Katabolismus mittels an sich bekannter mikrobiologischer Züchtungsmethoden erzielt wird.

25

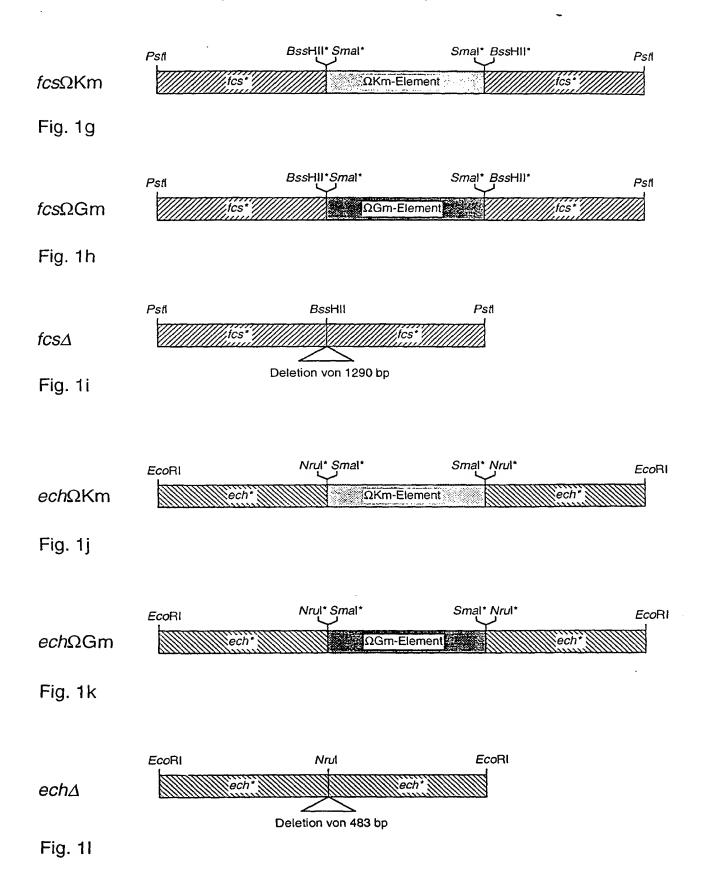
30

14. Verfahren zur Herstellung eines Organismus nach einem der Ansprüche 1 bis 5 oder 10 bis 11, dadurch gekennzeichnet, daß die Veränderung des Eugenolund/oder Ferulasäure-Katabolismus und/oder die Inaktivierung der entsprechenden Gene mittels gentechnischer Methoden erzielt wird.

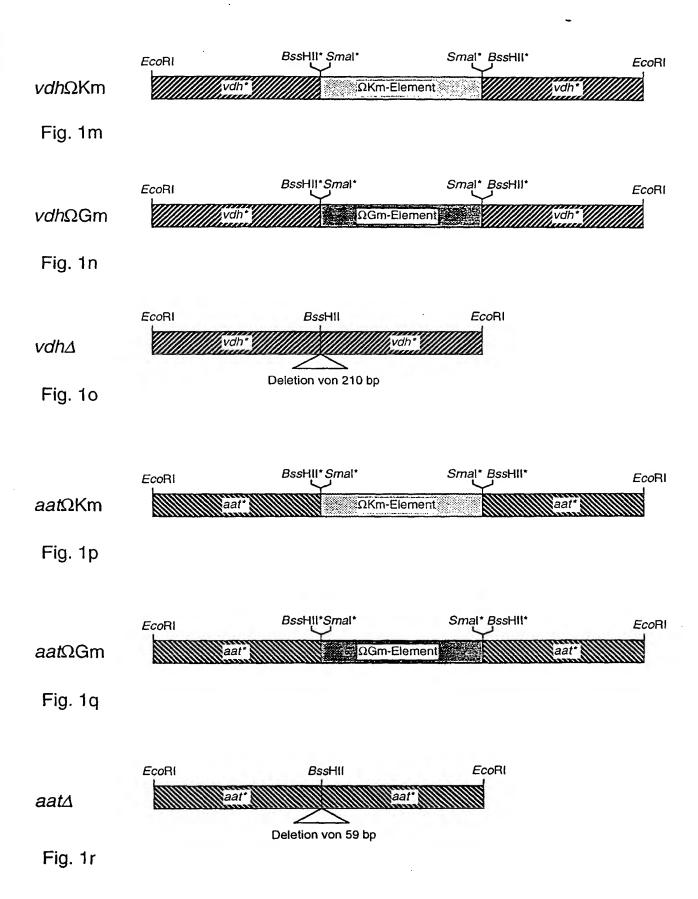
- 15. Verwendung der Organismen nach einem der Ansprüche 1 bis 5 oder 10 bis11 zur Herstellung von Coniferylalkohol, Coniferylaldehyd, Ferulasäure,Vanillin und/oder Vanillinsäure.
- 5 16. Verwendung von Gen-Strukturen nach einem der Ansprüche 6 bis 8 oder eines Vektors nach Anspruch 9 zur Herstellung transformierter und/oder mutagenisierter Organismen.



•
•



		•
		•
		•
		•



	-
,	
· ·	
•	

# Sequenzen

CTGCAGCCAG	GGCTGAA	AAG GAG	GGATTCA	GTGAG	GGTCAT	GAAGGGAC	GG GAC	GGCGCCT	60
GGCTCCAATT	GCTCGAT	GGC GCC	GCGATTG	AGTG	rcttgg	GCGCGGT	CTT GGA	GAGTTCG	120
GCTAGGGAGA	TTTAAAT	GCT GGC	CATGGTG	GCGG	CCCTG	ATGGGTT	GA TGA	TTTTCTG	180
CATTCTGCAT	CATGAAA	TTC ATG	AAATCAT	CACT	TTTCGG	GGGGTGG	STG CAC	GGGATTG	240
AAGGTTGCTA	GGAGAGI	GCA TTG	CTCGTAA	GCCC	AGGAAG	CACGCGG	STT TCA	GGATGGT	300
GCATGGAAAT	GGCATGA	AGCT TTG	CTGGATA	TGAT	TAGAGA	CATTAACT	TAT TAT	GGCGGAA	360
TGGAAGCACG	ATTCCTC	CGCC CGG	TAGAGCG	GTAA	CCGCGA	CATTCAG	GAC CGT	AAAAAGG	420
AAAGAGCATG Met 1	Gln Leu								<del>.</del>
GGT ATC GG Gly Ile G1	y Ala Gl			Val L			Gly Al		520
GTG ATT GG Val Ile Gl									568
GTT CAG GC Val Gln Al							Ala Il		616
ACAGCAAGCG	AACCGG	AATT GCO	CAGCTGGG	GCGC	CCTCTG	GTAAGGT	TGG GAA	GCCCTGC	676
AAAGTAAACT	GGATGG	CTTT CT	rgccgcc <i>i</i>	A AGGA	TCTGAT	GGCGCAG	GGG ATC	CAAGATCT	736
GATCAAGAGA	A CAGGAT	GAGG ATO	CGTTTCGC			A CAA GA J Gln As			790
GCA GGT TO Ala Gly Se 10									838
GCA CAA CA Ala Gln Gl 25									886
GCG CAG GG Ala Gln G							Ser G		934
CTG AAT G									982

, ·
0.27
77,
6.0
2
1

ACG GGC GTT CCT TGC GCA GCT GTG CTC GAC GTT GTC ACT GAA GCG GGA 1030 Thr Gly Val Pro Cys Ala Ala Val Leu Asp Val Val Thr Glu Ala Gly AGG GAC TGG CTG CTA TTG GGC GAA GTG CCG GGG CAG GAT CTC CTG TCA 1078 Arg Asp Trp Leu Leu Gly Glu Val Pro Gly Gln Asp Leu Leu Ser TCT CAC CTT GCT CCT GCC GAG AAA GTA TCC ATC ATG GCT GAT GCA ATG 1126 Ser His Leu Ala Pro Ala Glu Lys Val Ser Ile Met Ala Asp Ala Met 110 CGG CGG CTG CAT ACG CTT GAT CCG GCT ACC TGC CCA TTC GAC CAC CAA 1174 Arg Arg Leu His Thr Leu Asp Pro Ala Thr Cys Pro Phe Asp His Gln 125 GCG AAA CAT CGC ATC GAG CGA GCA CGT ACT CGG ATG GAA GCC GGT CTT 1222 Ala Lys His Arg Ile Glu Arg Ala Arg Thr Arg Met Glu Ala Gly Leu GTC GAT CAG GAT GAT CTG GAC GAA GAG CAT CAG GGG CTC GCG CCA GCC 1270 Val Asp Gln Asp Asp Leu Asp Glu Glu His Gln Gly Leu Ala Pro Ala 155 GAA CTG TTC GCC AGG CTC AAG GCG CGC ATG CCC GAC GGC GAG GAT CTC 1318 Glu Leu Phe Ala Arg Leu Lys Ala Arg Met Pro Asp Gly Glu Asp Leu 175 170 GTC GTG ACC CAT GGC GAT GCC TGC TTG CCG AAT ATC ATG GTG GAA AAT 1366 Val Val Thr His Gly Asp Ala Cys Leu Pro Asn Ile Met Val Glu Asn 190 195 GGC CGC TTT TCT GGA TTC ATC GAC TGT GGC CGG CTG GGT GTG GCG GAC 1414 Gly Arg Phe Ser Gly Phe Ile Asp Cys Gly Arg Leu Gly Val Ala Asp 205 CGC TAT CAG GAC ATA GCG TTG GCT ACC CGT GAT ATT GCT GAA GAG CTT 1462 Arg Tyr Gln Asp Ile Ala Leu Ala Thr Arg Asp Ile Ala Glu Glu Leu 225 220 GGC GGC GAA TGG GCT GAC CGC TTC CTC GTG CTT TAC GGT ATC GCC GCT 1510 Gly Gly Glu Trp Ala Asp Arg Phe Leu Val Leu Tyr Gly Ile Ala Ala 235 CCC GAT TCG CAG CGC ATC GCC TTC TAT CGC CTT CTT GAC GAG TTC TTC 1558 Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu Asp Glu Phe Phe 255 250 TGAGCGGGAC TCTGGGGTTC GAAATGACCG ACCAAGCGAC GCCCTG GCC GCG GTG 1613 Ala Ala Val 225 ATT GCA TTC ATG TGT GCT GAG GAG TCA CGT TGG ATC AAC GGC ATA AAT 1661

Ile Ala Phe Met Cys Ala Glu Glu Ser Arg Trp Ile Asn Gly Ile Asn

230

•
20
•
•

WO 00/26355 PCT/EP99/07952

ATT CCA GTG GAC GGA GGT TTG GCA TCG ACC TAC GTG TAA GTTCGTGGAC  Ile Pro Val Asp Gly Gly Leu Ala Ser Thr Tyr Val  245 250 255	1710
GCCCTTTGCA CGCGCACTAT ATCTCTATGC AGCAGCTGAA AGCAGCTTTG GTTTTGATCG	1770
GAGGTAGCGG GCGGAAAGGT GCAGAATGTC TAAATAATAA AGGATTCTTG TGAAGCTTTA	1830
GTTGTCCGTA AACGAAAATA AAAATAAAGA GGAATGATAT GAAAGCAAGT AGATCAGTCT	1,890
GCACTTTCAA AATAGCTACC CTGGCAGGCG CCATTTATGC AGCGCTGCCA ATGTCAGCTG	1950
CAAACTCGAT GCAGCTGGAT GTAGGTAGCT CGGATTGGAC GGTGCGTTGG GGACAACACC	2010
CTCAAGTATA GCCTTGCCTC TCGCCTGAAT GAGCAAGACT CAAGTCTGAC AAATGCGCCG	2070
ACTGTCAATG GTTATATCCG GATATTCAAA GTCAGGGTGA TCGTAACTTT GACCGGGGGC	2130
TTGGTATCCA ATCGTCTCGA TATTCTGGCT GCAG	2164
FIG. 2a:	

		•
		•
0.00		
		•

-	
CTGCAGCCAG GGCTGAAAAG GAGGGATTCA GTGAGGTCAT GAAGGGAGGG GACGGCGCCT	60
GGCTCCAATT GCTCGATGGC GCCGCGATTG AGTGTCTTGG GCGCGGTCTT GGAGAGTTCG	120
GCTAGGGAGA TAAATTTGCT GGCCATGGTG GCGGCCCCTG ATGGGTTGGA TGATTTTCTG	180
CATTCTGCAT CATGAAATTC ATGAAATCAT CACTTTTCGG GGGGTGGGTG CACGGGATTG	240
AAGGTTGCTA GGAGAGTGCA TTGCTCGTAA GCCCAGGAAG CACGCGGGTT TCAGGATGGT	300
GCATGGAAAT GGCATGAGCT TTGCTGGATA TGATTAGAGA CATTAACTAT TTTGGCGGAA	360
TGGAAGCACG ATTCCTCGCC CGGTAGAGCG GTAACCGCGA CATTCAGGAC CGTAAAAAGG	420
AAAGAGCATG CAA CTG ACC AAC AAG AAA ATC GTC GTC ACC GGA GTG TCC TCC  Met Gln Leu Thr Asn Lys Lys Ile Val Val Thr Gly Val Ser Ser  1 5 10 15	472
GGT ATC GGT GCC GAA ACT GCC CGC GTT CTG CGC TCT CAC GGC GCC ACA Gly Ile Gly Ala Glu Thr Ala Arg Val Leu Arg Ser His Gly Ala Thr 20 25 30	520
GTG ATT GGC GTA GAT CGC AAC ATG CCG AGC CTG ACT CTG GAT GCT TTC Val Ile Gly Val Asp Arg Asn Met Pro Ser Leu Thr Leu Asp Ala Phe 35 40 45	568
GTT CAG GCT GAC CTG AGC CAT CCT GAGGGGAGAG GCGGTTTGCG TATTGGGCGC Val Gln Ala Asp Leu Ser His Pro 50 55	622
ATGCATAAAA ACTGTTGTAA TTCATTAAGC ATTCTGCCGA CATGGAAGCC ATCACAAACG	682
GCATGATGAA CCTGAATCGC CAGCGGCATC AGCACCTTGT CGCCTTGCGT ATAATATTTG	742
CCCATGGACG CACACCGTGG AAACGGATGA AGGCACGAAC CCAGTTGACA TAAGCCTGTT	802
CGGTTCGTAA ACTGTAATGC AAGTAGCGTA TGCGCTCACG CAACTGGTCC AGAACCTTGA	862
CCGAACGCAG CGGTGGTAAC GGCGCAGTGG CGGTTTTCAT GGCTTGTTAT GACTGTTTTT	922
TTGTACAGTC TATGCCTCGG GCATCCAAGC AGCAAGCGCG TTACGCCGTG GGTCGATGTT	982
TGATGTTATG GAGCAGCAAC G ATG TTA CGC AGC AGC AAC GAT GTT ACG CAG  Met Leu Arg Ser Ser Asn Asp Val Thr Gln  1 5 10	1033
CAG GGC AGT CGC CCT AAA ACA AAG TTA GGT GGC TCA AGT ATG GGC ATC Gln Gly Ser Arg Pro Lys Thr Lys Leu Gly Gly Ser Ser Met Gly Ile  15 20 25	1081
ATT CGC ACA TGT AGG CTC GGC CCT GAC CAA GTC AAA TCC ATG CGG GCT Ile Arg Thr Cys Arg Leu Gly Pro Asp Gln Val Lys Ser Met Arg Ala 30 35 40	1129
GCT CTT GAT CTT TTC GGT CGT GAG TTC GGA GAC GTA GCC ACC TAC TCC Ala Leu Asp Leu Phe Gly Arg Glu Phe Gly Asp Val Ala Thr Tyr Ser 45 50 55	1177

.

				GAC Asp												1225
				CTT Leu												1273
				GTT Val 95												1321
				CTC Leu												1369
				ATC Ile												1417
				TAC Tyr												1465
	Leu			AAG Lys												1513
				ACC Thr 175	Ala			CAA	TTCG	TTC .	AAGC	CGAG	AT C	GGCT	TCCCT	1567
1		la P					lu G					le A			TA AAT le Asn	1616
		Val		GGA Gly			Ala						GTT	CGTG	GAC	1665
GCC	CTTT	'GCA	CGCG	CACI	'AT A	тстс	TATG	C AG	CAGC	TGAA	AGC	AGCT	TTG	GTTT	TGATCG	1725
GAC	GTAG	CGG	GCGG	DAAAG	GT G	CAGA	ATGT	C TA	ATAA	AATA	AGG	ATTC	TTG	TGAA	GCTTTA	1785
GTT	GTCC	GTA	AACC	AAAA	ATA A	TAAA	'AAAG	A GG	AATG	TATA	GAA	AGCA	AGT	AGAT	CAGTCT	1845
GCA	CTTI	CAA	AATA	AGCT	ACC C	TGGC	AGGC	G CC	TTTA	'ATGC	AGC	GCTG	CCA	ATGT	CAGCTG	1905
CA	AACTO	CGAT	GCAC	CTG	SAT G	STAGO	TAGO	T CO	GATI	'GGAC	GGT	GCGI	TGG	GGAC	CAACACC	1965
СТ	CAAGT	ATA	GCC	rtgco	CTC T	CGCC	TGAF	AT GA	AGCAP	AGACT	CAF	GTCI	GAC	PAAA	GCGCCG	2025
AC'	rgtc <i>i</i>	<b>A</b> ATG	GTT	TATC	CCG (	FATA	TCA	AA G1	CAGO	GTG <i>I</i>	TCC	CAAT	CTTT	GACC	CGGGGGC	2085
	GGTA:		ATC	GTCT(	CGA 1	TTAT	TGG	CT GO	CAG							2119

		•
		•
		•
		2-12
		•

CTGCAGCCAG GGCTGAAAAG GAGGGATTCA GTGAGGTCAT GAAGGGAGGG GACGGCGCCT	60
GGCTCCAATT GCTCGATGGC GCCGCGATTG AGTGTCTTGG GCGCGGTCTT GGAGAGTTCG	120
GCTAGGGAGA TAAATTTGCT GGCCATGGTG GCGGCCCCTG ATGGGTTGGA TGATTTTCTG	180
CATTCTGCAT CATGAAATTC ATGAAATCAT CACTTTTCGG GGGGTGGGTG CACGGGATTG	240
AAGGTTGCTA GGAGAGTGCA TTGCTCGTAA GCCCAGGAAG CACGCGGGTT TCAGGATGGT	,300
GCATGGAAAT GGCATGAGCT TTGCTGGATA TGATTAGAGA CATTAACTAT TTTGGCGGAA	360
TGGAAGCACG ATTCCTCGCC CGGTAGAGCG GTAACCGCGA CATTCAGGAC CGTAAAAAGG	420
AAAGAGCATG CAA CTG ACC AAC AAG AAA ATC GTC GTC ACC GGA GTG TCC TCC  Met Gln Leu Thr Asn Lys Lys Ile Val Val Thr Gly Val Ser Ser  1 5 10 15	472
GGT ATC GGT GCC GAA ACT GCC CGC GTT CTG CGC TCT CAC GGC GCC ACA Gly Ile Gly Ala Glu Thr Ala Arg Val Leu Arg Ser His Gly Ala Thr 20 25 30	520
GTG ATT GGC GTA GAT CGC AAC ATG CCG AGC CTG ACT CTG GAT GCT TTC Val Ile Gly Val Asp Arg Asn Met Pro Ser Leu Thr Leu Asp Ala Phe 35	568
GTT CAG GCT GAC CTG AGC CAT CCT GAA GGC ATC GATC AAC GGC ATA AAT Val Gln Ala Asp Leu Ser His Pro Glu Gly Ile Asn Gly Ile Asn 50 55 58 240	617
ATT CCA GTG GAC GGA GGT TTG GCA TCG ACC TAC GTG TAA GTTCGTGGAC Ile Pro Val Asp Gly Gly Leu Ala Ser Thr Tyr Val 245	666
GCCCTTTGCA CGCGCACTAT ATCTCTATGC AGCAGCTGAA AGCAGCTTTG GTTTTGATCG	726
GAGGTAGCGG GCGGAAAGGT GCAGAATGTC TAAATAATAA AGGATTCTTG TGAAGCTTTA	786
GTTGTCCGTA AACGAAAATA AAAATAAAGA GGAATGATAT GAAAGCAAGT AGATCAGTCT	846
GCACTTTCAA AATAGCTACC CTGGCAGGCG CCATTTATGC AGCGCTGCCA ATGTCAGCTG	906
CAAACTCGAT GCAGCTGGAT GTAGGTAGCT CGGATTGGAC GGTGCGTTGG GGACAACACC	966
CTCAAGTATA GCCTTGCCTC TCGCCTGAAT GAGCAAGACT CAAGTCTGAC AAATGCGCCG	1026
ACTGTCAATG GTTATATCCG GATATTCAAA GTCAGGGTGA TCGTAACTTT GACCGGGGGC	1086
TTGGTATCCA ATCGTCTCGA TATTCTGGCT GCAG	1120
FIG. 2c:	

•
•
•
•

GAATTO	CCGC	G TA	ATCG	CCG	G TT	CTAT	CAGC	GGG	CCGC	TTT	CGAA	AGTC	AT G	GTGT	TAGCC	60
GGTAG	GGTC	т т	TTTC	rtgg	C CA	TGCT	TGTT	GCC	TGAA	CCT	TCGT	TGAC	AT A	.GGGC	AGAGG	120
TGCGT	TTGC	CC GC	CTTC	GCTT	C GC	GATG.	AACC	GCA	TCGA	GAT.	GCTG	AGGT	CA G	GATT	TTTCC	180
TTAAC	TCGC	CG TA	AAGC	ATTC	T GT	CATT	TTTT	TGG	TGGC	TTT	GAAC	AGCC	TG A	TGAA	AGGTG	240
GTCTC	GCCC	CT T	rgag	GCCG	A TT	CTTG	GGCG	CTT	'GGCG	GCG	TCGA	AGCG	AT G	CTCC	ACTAC	.300
CGATT	AAG <i>P</i>	A TA	ATTA	TAAA	A AG	GAAA	CCGC	: ATG	GTTT	CTT	ATGT	'GAAT	TT G	TCTG	GCATA	360
CTCCA	GCTC	CA A	GGGC.	TTAA	т тт	GGGC	TTAT	· GGC	TGAG	CAG	TTGC	CTCT	A TA	TGGI	TATTC	420
AGAAT	AAC	AA T'	TGAC	TCCT	C AG	GAGG	TCAG	CG				CTT Leu				473
GGT G Gly A																521
AAG A Lys L																569
AGT A Ser A 40																617
ATT G Ile A																665
ACA C																713
CGC G Arg G																761
TTT ( Phe F																809
GGG ( Gly V 120																857
CTG ( Leu <i>l</i>					Ala					Ala					Ser	905
GAG (									Leu					Ala		953

•

.

į

	Phe													GAA Glu		1001
GGT Gly					_									ACC Thr		1049
														AAC Asn		1.097
														GTT Val 230		1145
														GTG Val		1193
														CTG Leu		1241
	GAA Glu 265	GGG	ACAG	CAA (	GCGA	ACCG	GA A'	TTGC	CAGC'	r GG	GGCG	CCCT	CTG	GTAA	GGT	1297
TGG	GAAG	ccc '	TGCA	AAGT.	AA AA	CTGG	ATGG	C TT	TCTT	GCCG	CCA	AGGA'	TCT (	GATG	GCGCAG	1357
GGG	ATCA.	AGA '	TCTG	ATCA	AG A	GACA	GGAT	G AG	GATC	GTTT	CGC			GAA Glu		1412
														CTA Leu		1460
					Gln					Cys				GCC Ala 35	Val	1508
				Ala					Val						GAC Asp	1556
			Ala					Gln					Arg		TCG Ser	1604
		Ala					Pro					Leu			GTC Val	1652

	-		
			•
			•
			•

									TTG Leu							1700
									GCC Ala 110							1748
									CTT Leu							1796
									GAG Glu							1844
									CTG Leu							1892
									CTC Leu							1940
									GAT Asp 190							1988
									TTC Phe							2036
			Asp					Ile							ATT	2084
GCT Ala	GAA Glu 230	Glu	CTT Leu	GGC Gly	GGC	GAA Glu 235	TGG Trp	GCT Ala	GAC Asp	CGC Arg	TTC Phe 240	Leu	GTG Val	CTT Leu	TAC	2132
	Ile		Ala	Pro		Ser	Gln	Arg	Ile		Phe				CTT Leu 260	2180
			TTC Phe	+	GCG	GGAC	TCT	GGGG	TTCG	AA A	TGAC	CGAC	C AA	GCGP	CGCC	2235
CGC	His		Lys					Glr					Glu		AAC Asn	2283
		Met					Gl;					Gly			C TCT 1 Ser	2331

		4
		•
	<b>X</b>	
		·
		-

	TCA ACG GA Ser Thr Gl 48		AACCGTTGGT i	AGTGGTTTTG (	GACGGGCCCA	2385
GGAGCATGCG	CTTCTGGGCC	CGTTTCTTGA	GTATTCATTG	GATAGTCACG	CGTGGTAGCT	2445
TCGAGCCTGC	ACAGCTGATG	AGCACCCTGG	AAGGCGCGCT	GTACGCGGAC	GACTGGGTTC	2505
ATCTTCGCCA	TTCATGACGG	AACTCCGTTC	CCCAGTACCG	CGATGACTAT	TTTGCCTCTT	2565
CCGATGTCCG	ATTCCACGCC	GCCTGACGCT	AAGCGGGGGC	GGGGGCGCCC	GCATCCCAGC	2625
CCAGACAGCA	ACAAATGAGT	AGGCTCTTGG	ATGCCGCGGC	GGCTGAGATT	GGTAACGGCA	2685
ATTTCGTCAA	TGTGACGATG	GATTCGATTG	CCCGTGCTGC	CGGCGTCTCA	AAAAAAACGC	2745
TGTACGTCTT	GGTGGCGAGC	AAGGAAGAAC	TCATTTCCCG	GTTAGTGGCT	CGAGACATGT	2805
CCAACCTTGA	GGAATTC					2822
FIG. 2d:						

•
¥
•
<u> </u>

GAATTCCGCG TA	TCGCCCGG TT	CTATCAGC GG	GCCGCTTT	CGAAAGTCAT	GGTGTTAGCC	60
GGTAGGGTCT TT	TTCTTGGC CA	rgcttgtt gc	CTGAACCT	TCGTTGACAT	AGGGCAGAGG	120
TGCGTTTGCC GC	TTCGCTTC GC	GATGAACC GC	ATCGAGAT	GCTGAGGTCA	GGATTTTTCC	180
TTAACTCGCG TA	AGCATTCT GT	CATTTTTT TG	GTGGCTTT	GAACAGCCTG	ATGAAAGGTG	240
GTCTCGCCCT TT	GAGGCCGA TT	CTTGGGCG CT	TGGCGGCG	TCGAAGCGAT	GCTCCACTAC	300
CGATTAAGAT AA	ттаааата ас	GAAACCGC AT	GGTTTCTT	ATGTGAATTT	GTCTGGCATA	360
CTCCAGCTCA AG	GGCAATTT TT	GGGCTATT GG	CTGAGCAG	TTGCCTCTAT	ATGGTTATTC	420
AGAATAACAA TT	GACTCCTC AG	GAGGTCAG CG		ATT CTT GGT Ile Leu Gly	Leu Asn	473
GGT GCC CCG G Gly Ala Pro V 10						521
AAG AAG GCG C Lys Lys Ala H 25						569
AGT AGG CTG C Ser Arg Leu A 40						617
ATT GCC GAC ( Ile Ala Asp A						665
ACA CTG CTT C			Val Ala		s Asp Ser	713
CGC GAG CAC ( Arg Glu His '						761
TTT CCA GGG (Phe Pro Gly 1	GCG GAG GCA Ala Glu Ala	CGC GTT GAC Arg Val Glu 110	S TTT CAG Phe Gln	CCG CTG GG Pro Leu Gl 115	T GTC GTT y Val Val	809
GGG GTC ATT Gly Val Ile 120				Leu Ala Ph		857
CTG GCC GGC Leu Ala Gly						905
GAG CTT ACC Glu Leu Thr			u Leu Ala		e Ala Arg	953

•
•
•
•
į

TAC TTC GAT GAA ACT GAG CTG ACT ACA GTG CTG GGC GAC GCT GAA GTC Tyr Phe Asp Glu Thr Glu Leu Thr Thr Val Leu Gly Asp Ala Glu Val 170 175 180	1001
GGT GCG CTG TTC AGT GCT CAG CCT TTC GAT CAT CTG ATC TTC ACC GGC Gly Ala Leu Phe Ser Ala Gln Pro Phe Asp His Leu Ile Phe Thr Gly 185	1049
GGC ACT GCC GTG GCC AAG CAC ATC ATG CGT GCC GCG GCG GAT AAC CTA Gly Thr Ala Val Ala Lys His Ile Met Arg Ala Ala Ala Asp Asn Leu 200 205 210 215	1097
GTG CCC GTT ACC CTG GAA TTG GGT GGC AAA TCG CCG GTG ATC GTT TCC Val Pro Val Thr Leu Glu Leu Gly Gly Lys Ser Pro Val Ile Val Ser 220 225 230	1145
CGC AGT GCA GAT ATG GCG GAC GTT GCA CAA CGG GTG TTG ACG GTG AAA Arg Ser Ala Asp Met Ala Asp Val Ala Gln Arg Val Leu Thr Val Lys 235 240 245	1193
ACC TTC AAT GCC GGG CAA ATC TGT CTG GCA CCG GAC TAT GTG CTG GGG Thr Phe Asn Ala Gly Gln Ile Cys Leu Ala Pro Asp Tyr Val Leu 250 262	1241
GAGAGGCGGT TTGCGTATTG GGCGCATGCA TAAAAACTGT TGTAATTCAT TAAGCATTCT	1301
GCCGACATGG AAGCCATCAC AAACGGCATG ATGAACCTGA ATCGCCAGCG GCATCAGCAC	1361
CTTGTCGCCT TGCGTATAAT ATTTGCCCAT GGACGCACAC CGTGGAAACG GATGAAGGCA	1421
CGAACCCAGT TGACATAAGC CTGTTCGGTT CGTAAACTGT AATGCAAGTA GCGTATGCGC	1481
TCACGCAACT GGTCCAGAAC CTTGACCGAA CGCAGCGGTG GTAACGGCGC AGTGGCGGTT	1541
TTCATGGCTT GTTATGACTG TTTTTTTGTA CAGTCTATGC CTCGGGCATC CAAGCAGCAA	1601
GCGCGTTACG CCGTGGGTCG ATGTTTGATG TTATGGAGCA GCAACG ATG TTA CGC Met Leu Arg 1	1656
AGC AGC AAC GAT GTT ACG CAG CAG GGC AGT CGC CCT AAA ACA AAG TTA Ser Ser Asn Asp Val Thr Gln Gln Gly Ser Arg Pro Lys Thr Lys Leu 5 10 15	1704
GGT GGC TCA AGT ATG GGC ATC ATT CGC ACA TGT AGG CTC GGC CCT GAC Gly Gly Ser Ser Met Gly Ile Ile Arg Thr Cys Arg Leu Gly Pro Asp 20 25 30 35	1752
CAA GTC AAA TCC ATG CGG GCT GCT CTT GAT CTT TTC GGT CGT GAG TTC Gln Val Lys Ser Met Arg Ala Ala Leu Asp Leu Phe Gly Arg Glu Phe 40 45 50	1800
GGA GAC GTA GCC ACC TAC TCC CAA CAT CAG CCG GAC TCC GAT TAC CTC Gly Asp Val Ala Thr Tyr Ser Gln His Gln Pro Asp Ser Asp Tyr Leu 55 60 65	1848
GGG AAC TTG CTC CGT AGT AAG ACA TTC ATC GCG CTT GCT GCC TTC GAC Gly Asn Leu Leu Arg Ser Lys Thr Phe Ile Ala Leu Ala Ala Phe Asp 70 75 80	1896

	•
	*

CAA GAA GCG GTT GTT GGC GCT CTC GCG GCT TAC GTT CTG CCC AGG TTT 1944 Gln Glu Ala Val Val Gly Ala Leu Ala Ala Tyr Val Leu Pro Arg Phe GAG CAG CCG CGT AGT GAG ATC TAT ATC TAT GAT CTC GCA GTC TCC GGC 1992 Glu Gln Pro Arg Ser Glu Ile Tyr Ile Tyr Asp Leu Ala Val Ser Gly 105 110 GAG CAC CGG AGG CAG GGC ATT GCC ACC GCG CTC ATC AAT CTC CTC AAG 2040 Glu His Arg Arg Gln Gly Ile Ala Thr Ala Leu Ile Asn Leu Lys 120 CAT GAG GCC AAC GCG CTT GGT GCT TAT GTG ATC TAC GTG CAA GCA GAT 2088 His Glu Ala Asn Ala Leu Gly Ala Tyr Val Ile Tyr Val Gln Ala Asp 140 TAC GGT GAC GAT CCC GCA GTG GCT CTC TAT ACA AAG TTG GGC ATA CGG 2136 Tyr Gly Asp Asp Pro Ala Val Ala Leu Tyr Thr Lys Leu Gly Ile Arg 155 150 GAA GAA GTG ATG CAC TTT GAT ATC GAC CCA AGT ACC GCC ACC TAA CAA 2184 Glu Glu Val Met His Phe Asp Ile Asp Pro Ser Thr Ala Thr 170 TTCGTTCAAG CCGAGATCGG CTTCCCTG CAA AGT CCT GTG GGT GAG TCG AAC 2236 Gln Ser Pro Val Gly Glu Ser Asn TTG GCG ATG CGC GCA CCC TAC GGA GAA GCG ATC CAC GGA CTG CTC TCT 2284 Leu Ala Met Arq Ala Pro Tyr Gly Glu Ala Ile His Gly Leu Leu Ser 460 465 GTC CTC CTT TCA ACG GAG TGT TAG AACCGTTGGT AGTGGTTTTG GACGGGCCCA 2338 Val Leu Leu Ser Thr Glu Cys 475 480 481 GGAGCATGCG CTTCTGGGCC CGTTTCTTGA GTATTCATTG GATAGTCACG CGTGGTAGCT 2398 TCGAGCCTGC ACAGCTGATG AGCACCCTGG AAGGCGCGCT GTACGCGGAC GACTGGGTTC 2458 ATCTTCGCCA TTCATGACGG AACTCCGTTC CCCAGTACCG CGATGACTAT TTTGCCTCTT 2518 CCGATGTCCG ATTCCACGCC GCCTGACGCT AAGCGGGGGC GGGGGCGCCC GCATCCCAGC 2578 CCAGACAGCA ACAAATGAGT AGGCTCTTGG ATGCCGCGGC GGCTGAGATT GGTAACGGCA 2638 ATTTCGTCAA TGTGACGATG GATTCGATTG CCCGTGCTGC CGGCGTCTCA AAAAAAACGC 2698 TGTACGTCTT GGTGGCGAGC AAGGAAGAAC TCATTTCCCG GTTAGTGGCT CGAGACATGT 2758 CCAACCTTGA GGAATTC 2775

FIG. 2e:

•
•
į

GAATTCCGCG TATCGCCCG	G TTCTATCAGC	GGGCCGCTTT	CGAAAGTCAT GGTGTTAGCC	60
GGTAGGGTCT TTTTCTTGG	C CATGCTTGTT	GCCTGAACCT	TCGTTGACAT AGGGCAGAGG	120
TGCGTTTGCC GCTTCGCTT	C GCGATGAACC	GCATCGAGAT	GCTGAGGTCA GGATTTTTCC	180
TTAACTCGCG TAAGCATTC	CT GTCATTTTTT	TGGTGGCTTT	GAACAGCCTG ATGAAAGGTG	240
GTCTCGCCCT TTGAGGCCC	SA TTCTTGGGCG	CTTGGCGGCG	TCGAAGCGAT GCTCCACTAC	300
CGATTAAGAT AATTAAAA	A AGGAAACCGC	ATGGTTTCTT	ATGTGAATTT GTCTGGCATA	360
CTCCAGCTCA AGGGCAATT	TT TTGGGCTATT	GGCTGAGCAG	TTGCCTCTAT ATGGTTATTC	420
AGAATAACAA TTGACTCC	rc aggaggtcag		ATT CTT GGT TTG AAT Ile Leu Gly Leu Asn 5	473
GGT GCC CCG GTC GGA Gly Ala Pro Val Gly 10	•			521
AAG AAG GCG CAC CTG Lys Lys Ala His Leu 25				569
AGT AGG CTG GAT CGT Ser Arg Leu Asp Arg 40				617
ATT GCC GAC GCG GTT Ile Ala Asp Ala Val 60	Ser Ala Asp			665
ACA CTG CTT TGC GAC Thr Leu Leu Cys Asp 75				713
CGC GAG CAC GTG GCC Arg Glu His Val Ala 90			CAT CAC AAG GCG ATG His His Lys Ala Met 100	761
			CCG CTG GGT GTC GTT Pro Leu Gly Val Val 115	809
			CTG GCC TTT GGG CCG Leu Ala Phe Gly Pro 135	857
	Ala Ala Gly		ATG CTC AAG CCG TCC Met Leu Lys Pro Ser 150	905
			GAG CTA ATT GCT CGT Glu Leu Ile Ala Arg 165	953

		٠
		è
		•

	Phe													GAA Glu		1001
														ACC Thr		1049
														AAC Asn		1097
-														GTT Val 230		1145
														GTG Vàl		1193
	TTC Phe		Ala							•		Gly (		rcg / Ser /		1240
		Met												CTC Leu		1288
	Leu				Glu		TAG	AAC	CGTT	GGT :	AGTG	GTTT'	TG G.	ACGG	GCCCA	1342
GGA	GCAT	GCG	СТТС	TGGG	cc c	GTTT	CTTG.	A GT.	ATTC.	ATTG	GAT.	AGTC.	ACG	CGTG	GTAGCT	1402
TCG	AGCC	TGC	ACAG	CTGA	TG A	GCAC	CCTG	G AA	GGCG	CGCT	GTA	CGCG	GAC	GACT	GGGTTC	1462
ATC	TTCG	CCA	TTCA	TGAC	GG A	ACTC	CGTT	c cc	CAGT	ACCG	CGA	TGAC	TAT	TTTG	CCTCTT	1522
CCG	ATGT	CCG	ATTC	CACG	CC G	CCTG	ACGC	T AA	GCGG	GGGC	GGG	GGCG	CCC	GCAT	CCCAGC	1582
CCA	GACA	GCA	ACAA	ATGA	GT A	GGCT	CTTG	G AT	GCCG	CGGC	GGC	TGAG	TTA	GGTA	ACGGCA	1642
ATT	TCGT	CAA	TGTG	ACGA	TG G	ATTC	GATT	G CC	CGTG	CTGC	CGG	CGTC	TCA	AAAA	AAACGC	1702
TGT	'ACGT	CTT	GGTG	GCGA	GC A	AGGA	AGAA	C TC	TTTA	CCCG	GTT	AGTG	GCT	CGAG	ACATGT	1762
CCA	ACCT	TGA	GGAA	TTC												1779

FIG. 2f:


•

CTGC	AGCC	GA G	CATC	GATT	g ag	CACT	TTAC	CCA	GCTG	CGC	TGGC	TGAC	CA I	TCAG	AATGG	60	
CCCG	CGGC	AC T	ATCC	AATC	AA T	ATCG	ATCT	TCG	GGCG	CCG	CGGG	CATC	AT C	cccg	CGGCG	120	
CTCG	ССТС	T TA	TCAA	TCTC	T AA	CTTG	ATAA	. AAA	CAGA	GCT	GTTC	TCCG	GT C	TTGG	TGGAT	180	
CAAG	GCCA	GT C	GCGG	AGAG	т ст	CGAA	GAGG	AGA	GTAC	AGT	GAAC	CCCG	AG T	CCAC	CATTGC	240	
AACC	GCAG	GC A	TCAT	CATG	C TC	TGCT	CAGO	CAC	GCTA	CCG	CAGI	GTGT	CG F	ATTGG	STCATC	300	
CTCC	GGTT	GA G	GTTA	.CGCA	A GA	CGCT	GGAG	GTA	lttg1	CCG					C GAG u Glu 5	356	
GCG Ala																404	
														GCA Ala		452	
														GTC Val		500	
														CGT Arg		548	
														GCA Ala		596	,
														GCT Ala 100		644	
														GGT Gly		692	?
			Gly					Ala					Phe	CAG Gln 132		740	)
ACA	GCAA	GCG	AACC	GGAA	TT G	CCAG	CTGG	G GC	GCCC	TCTG	GTA	AGGT	TGG	GAAG	CCCTGC	800	)
AAA	GTAA	ACT	GGAT	GGCT	тт с	TTGC	CGCC	A AG	GATC	TGAT	, GGC	GCAG	GGG	ATCA	AGATCT	860	)
GAT	CAAG	AGA	CAGG	ATGA	.GG A	TCGT	TTCG								G CAC u His	914	3
		Ser					Val					Gly			TGG Trp	962	2

		(•)
		•

.

\*

			ACA Thr													1010
			CGC Arg													1058
			CTG Leu 60													1106
			CCT Pro													1154
			CTG Leu													1202
			GCT Ala													1250
			CAT His													1298
GCG Ala	AAA Lys	CAT	CGC Arg 140	Ile	GAG Glu	CGA Arg	GCA Ala	CGT Arg 145	ACT Thr	CGG Arg	ATG Met	GAA Glu	GCC Ala 150	GGT Gly	CTT Leu	1346
			Asp												GCC Ala	1394
		Phe					Ala					Gly			CTC Leu	1442
	Val			Gly		Ala	Cys	Leu		Asn	Ile				AAT Asn 200	1490
					Phe					Arg					GAC Asp	153
				o Ile					Arg					Glu	CTT Leu	158
GG( Gly	GGG Gly	GAA y Glu 23!	ı Trp	G GCT	GAC Asp	CGC Arg	TTC Phe 240	e Leu	GTC ı Val	CTI Leu	TAC 1 Ty1	GGT Gly 245	/ Ile	GCC Ala	GCT Ala	163

	_
•	
•	
•	

CCC GAT TCG CAG CGC ATC GCC TTC TAT CGC CTT CTT GAC GAG TTC TTC Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu Asp Glu Phe Phe 250 255 260 264	1682
TGAGCGGGAC TCTGGGGTTC GAAATGACCG ACCAAGCGAC GCCCCT GTT TTG CAA Val Leu Gln 563 565	1737
TGG CGG TCG GCG AAA GTT GAT GCG CTG TAT CGT GGT GAA GAT CAA TCC Trp Arg Ser Ala Lys Val Asp Ala Leu Tyr Arg Gly Glu Asp Gln Ser 570 575 580	1785
ATG CTG CGT GAC GAG GCC ACA CTG TGA GTTGGTCAGG GGGGGCTTAC  Met Leu Arg Asp Glu Ala Thr Leu  585 589	1832
TCGGCGTTTT CCGACACTGC GTTGGTTGCG GCAGTGCGCA CCCCCTGGAT TGATTGCGGG	1892
GGTGCCCTGT CGCTGGTGTC GCCTATCGAC TTAGGGGTAA AGGTCGCTCG CGAAGTTCTG	1952
ATGCGTGCGT CGCTTGAACC ACAAATGGTC GATAGCGTAC TCGCAGGCTC TATGGCTCAA	2012
GCAAGCTTTG ATGCTTACCT GCTCCCGCGG CACATTGGCT TGTACAGCGG TGTTCCCAAG	2072
TCGGTTCCGG CCTTGGGGGT GCAGCGCATT TGCGGCACAG GCTTCGAACT GCTTCGGCAG	2132
GCCGGCGAGC AGATTTCCCA AGGCGCTGAT CACGTGCTGT GTGTCGCGGG CTGCAG	2188
FIG. 2g:	

	-
<u>.</u>	
· · · · · · · · · · · · · · · · · · ·	
•	
•	

CTGCAGCCGA GCATCGATTG AGCACTTTAC CCAGCTGCGC TGGCTGACCA TTCAGAATGG	60
CCCGCGGCAC TATCCAATCT AAATCGATCT TCGGGCGCCG CGGGCATCAT GCCCGCGGCG	120
CTCGCCTCAT TTCAATCTCT AACTTGATAA AAACAGAGCT GTTCTCCGGT CTTGGTGGAT	180
CAAGGCCAGT CGCGGAGAGT CTCGAAGAGG AGAGTACAGT GAACGCCGAG TCCACATTGC	240
AACCGCAGGC ATCATCATGC TCTGCTCAGC CACGCTACCG CAGTGTGTCG ATTGGTCATC	300
CTCCGGTTGA GGTTACGCAA GACGCTGGAG GTATTGTCCG G ATG CGT TCT CTC GAG	356
Met Arg Ser Leu Glu 1 5	330
GCG CTT CTT CCC TTC CCG GGT CGA ATT CTT GAG CGT CTC GAG CAT TGG Ala Leu Leu Pro Phe Pro Gly Arg Ile Leu Glu Arg Leu Glu His Trp  10 15 20	404
GCT AAG ACC CGT CCA GAA CAA ACC TGC GTT GCT GCC AGG GCG GCA AAT Ala Lys Thr Arg Pro Glu Gln Thr Cys Val Ala Ala Arg Ala Ala Asn 25 30 35	452
GGG GAA TGG CGT CGT ATC AGC TAC GCG GAA ATG TTC CAC AAC GTC CGC Gly Glu Trp Arg Arg Ile Ser Tyr Ala Glu Met Phe His Asn Val Arg 40 45 50	500
GCC ATC GCA CAG AGC TTG CTT CCT TAC GGA CTA TCG GCA GAG CGT CCG Ala Ile Ala Gln Ser Leu Leu Pro Tyr Gly Leu Ser Ala Glu Arg Pro 55 60 65	548
CTG CTT ATC GTC TCT GGA AAT GAC CTG GAA CAT CTT CAG CTG GCA TTT Leu Leu Ile Val Ser Gly Asn Asp Leu Glu His Leu Gln Leu Ala Phe 70 75 80 85	596
GGG GCT ATG TAT GCG GGC ATT CCC TAT TGC CCG GTG TCT CCT GCT TAT Gly Ala Met Tyr Ala Gly Ile Pro Tyr Cys Pro Val Ser Pro Ala Tyr 90 95 100	644
TCA CTG CTG TCG CAA GAT TTG GCG AAG CTG CGT CAC ATC GTA GGT CTT Ser Leu Leu Ser Gln Asp Leu Ala Lys Leu Arg His Ile Val Gly Leu 105 110 115	692
CTG CAA CCG GGA CTG GTC TTT GCT GCC GAT GCA GCA CCT TTC CAG GGG Leu Gln Pro Gly Leu Val Phe Ala Ala Asp Ala Ala Pro Phe Gln 120 125 130 132	740
GAGAGGCGGT TTGCGTATTG GGCGCATGCA TAAAAACTGT TGTAATTCAT TAAGCATTCT	800
GCCGACATGG AAGCCATCAC AAACGGCATG ATGAACCTGA ATCGCCAGCG GCATCAGCAC	860
CTTGTCGCCT TGCGTATAAT ATTTGCCCAT GGACGCACAC CGTGGAAACG GATGAAGGCA	920
CGAACCCAGT TGACATAAGC CTGTTCGGTT CGTAAACTGT AATGCAAGTA GCGTATGCGC	980
TCACGCAACT GGTCCAGAAC CTTGACCGAA CGCAGCGGTG GTAACGGCGC AGTGGCGGTT	1040
TTCATGGCTT GTTATGACTG TTTTTTTGTA CAGTCTATGC CTCGGGCATC CAAGCAGCAA	1100

		-
		•
		•
		•

GCGCGTTACG CCGTGGGTCG ATGTTTGATG TTATGGAGCA GCAACG ATG TTA CGC  Met Leu Arg  1	1155
AGC AGC AAC GAT GTT ACG CAG CAG GGC AGT CGC CCT AAA ACA AAG TTA Ser Ser Asn Asp Val Thr Gln Gln Gly Ser Arg Pro Lys Thr Lys Leu 5	1203
GGT GGC TCA AGT ATG GGC ATC ATT CGC ACA TGT AGG CTC GGC CCT GAC Gly Gly Ser Ser Met Gly Ile Ile Arg Thr Cys Arg Leu Gly Pro Asp 20 25 30 35	1251
CAA GTC AAA TCC ATG CGG GCT GCT CTT GAT CTT TTC GGT CGT GAG TTC Gln Val Lys Ser Met Arg Ala Ala Leu Asp Leu Phe Gly Arg Glu Phe 40 45 50	1299
GGA GAC GTA GCC ACC TAC TCC CAA CAT CAG CCG GAC TCC GAT TAC CTC Gly Asp Val Ala Thr Tyr Ser Gln His Gln Pro Asp Ser Asp Tyr Leu 55 60 65	1347
GGG AAC TTG CTC CGT AGT AAG ACA TTC ATC GCG CTT GCT GCC TTC GAC Gly Asn Leu Arg Ser Lys Thr Phe Ile Ala Leu Ala Ala Phe Asp 70 75 80	1395
CAA GAA GCG GTT GTT GGC GCT CTC GCG GCT TAC GTT CTG CCC AGG TTT Glu Ala Val Val Gly Ala Leu Ala Ala Tyr Val Leu Pro Arg Phe 85 90 95	1443
GAG CAG CCG CGT AGT GAG ATC TAT ATC TAT GAT CTC GCA GTC TCC GGC Glu Gln Pro Arg Ser Glu Ile Tyr Ile Tyr Asp Leu Ala Val Ser Gly 100 105 110	1491
GAG CAC CGG AGG CAG GGC ATT GCC ACC GCG CTC ATC AAT CTC CTC AAG Glu His Arg Arg Gln Gly Ile Ala Thr Ala Leu Ile Asn Leu Leu Lys 120 125 130	1539
CAT GAG GCC AAC GCG CTT GGT GCT TAT GTG ATC TAC GTG CAA GCA GAT His Glu Ala Asn Ala Leu Gly Ala Tyr Val Ile Tyr Val Gln Ala Asp 135 140 145	1587
TAC GGT GAC GAT CCC GCA GTG GCT CTC TAT ACA AAG TTG GGC ATA CGG Tyr Gly Asp Asp Pro Ala Val Ala Leu Tyr Thr Lys Leu Gly Ile Arg 150 155 160	1635
GAA GAA GTG ATG CAC TTT GAT ATC GAC CCA AGT ACC GCC ACC TAA CAA Glu Glu Val Met His Phe Asp Ile Asp Pro Ser Thr Ala Thr 165 170 175 177	1683
TTCGTTCAAG CCGAGATCGG CTTCCCCT GTT TTG CAA TGG CGG TCG GCG AAA Val Leu Gln Trp Arg Ser Ala Lys 563 565 570	1735
GTT GAT GCG CTG TAT CGT GGT GAA GAT CAA TCC ATG CTG CGT GAC GAG Val Asp Ala Leu Tyr Arg Gly Glu Asp Gln Ser Met Leu Arg Asp Glu 575 580 585	1783

4.,		
		*

•

4

WO 00/26355 PCT/EP99/07952

GCC ACA CTG Ala Thr Leu 589		'CAGG GGGGGC	CTTAC TCGGCG	STTTT CCGACA	ACTGC	1835
GTTGGTTGCG	GCAGTGCGCA	CCCCTGGAT	TGATTGCGGG	GGTGCCCTGT	CGCTGGTGTC	1895
GCCTATCGAC	TTAGGGGTAA	AGGTCGCTCG	CGAAGTTCTG	ATGCGTGCGT	CGCTTGAACC	1955
ACAAATGGTC	GATAGCGTAC	TCGCAGGCTC	TATGGCTCAA	GCAAGCTTTG	ATGCTTACCT	2015
GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	TGTTCCCAAG	TCGGTTCCGG	CCTTGGGGGT	2075
GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	GCTTCGGCAG	GCCGGCGAGC	AGATTTCCCA	2135
AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGG	CTGCAG			2171
FIG. 2h:						

÷
•
,
•

CTGC	AGCC	GA G	CATC	GATT	G AG	CACT	TTAC	CCA	GCTG	CGC	TGGC	TGAC	CA I	TCAG	SAATGG	60
CCCG	CGGC	AC T	ATCC	AATC	T AA	ATCG	ATCT	TCG	GGCG	CCG	CGGG	CATO	CAT G	CCCG	CGGCG	120
CTCG	CCTC	T TA	TCAA	TCTC	T AA	CTTG	АТАА	. AAA	CAGA	GCT	GTTC	TCC	GT C	TTGG	TGGAT	180
CAAG	GCCA	GT C	GCGG	AGAG	т ст	CGAA	.GAGG	AGA	GTAC	CAGT	GAAC	GCCG	SAG T	CCAC	CATTGC	240
AACC	GCAG	GC A	TCAT	CATG	C TC	TGCT	CAGC	CAC	GCTA	ACCG	CAGT	GTGT	CG F	ATTGO	STCATC	300
CTCC	GGTT	GA G	GTTA	CGCA	A GA	CGCT	GGAG	GTP	TTGT	CCG					CC GAG eu Glu 5	356
			CCC Pro													404
			CGT Arg 25													452
			CGT Arg													500
			CAG Gln													548
			GTC Val													596
			TAT Tyr													644
			TCG Ser 105													692
			GGA Gly										Phe			740
	Val		CAA Gln 565	Trp										Arg	GGT Gly	788
			TCC Ser					Glu					GTT	GGTC	AGG	831
GGG	GGCT	TAC	TCGG	CGTT	тт с	CGAC	ACTG	C GT	TGGT	TGCG	GCA	GTGC	GCA	cccc	CTGGAT	89
TCλ	ምምርር	ccc	CCTC	CCCT	CT C	GCTG	CTCT	ר פר	СТАТ	ירהאר	ጥጥ አ	cccc	ממדי	AGGT	CGCTCG	95

		7
		•
		<b>,</b>
		4

WO 00/26355 PCT/EP99/07952

CGAAGTTCTG	ATGCGTGCGT	CGCTTGAACC	ACAAATGGTC	GATAGCGTAC	TCGCAGGCTC	1017
TATGGCTCAA	GCAAGCTTTG	ATGCTTACCT	GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	1077
TGTTCCCAAG	TCGGTTCCGG	CCTTGGGGGT	GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	1137
GCTTCGGCAG	GCCGGCGAGC	AGATTTCCCA	AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGG	1197
CTGCAG						1203
FIG. 2i:						

		,
		•
		*

GAATTCCCCT GGCGACGAAA GGGCGGCAGG CCGCATGGCC ACGGCTGGGC GGTAACTGAT	60
GCTTGCGTTA ATCGTTAACC GTTTGAAATT CCTTGCCAAA TTTCGGCGAG AGAATCATGC	120
GGGTACGCCT TTCCGTGCGC TTTGATCTGC GCTTCCGTGC CTTGAATCAG AAAAATAGTT	180
AATTGACAGA ACTATAGGTT CGCAGTAGCT TTTGCTCACC CACCAAATCC ACAGCACTGG	240
GGTGCACG ATG AAT AGC TAC GAT GGC CGT TGG TCT ACC GTT GAT GTG AAG  Met Asn Ser Tyr Asp Gly Arg Trp Ser Thr Val Asp Val Lys  1 5 10	290
GTT GAA GAA GGT ATC GCT TGG GTC ACG CTG AAC CGC CCG GAG AAG CGC Val Glu Glu Gly Ile Ala Trp Val Thr Leu Asn Arg Pro Glu Lys Arg 15 20 25 30	338
AAC GCA ATG AGC CCA ACT CTC AAT CGA GAG ATG GTC GAG GTT CTG GAG Asn Ala Met Ser Pro Thr Leu Asn Arg Glu Met Val Glu Val Leu Glu . 35	386
GTG CTG GAG CAG GAC GCA GAT GCT CGC GTG CTT GTT CTG ACT GGT GCA Val Leu Glu Gln Asp Ala Asp Ala Arg Val Leu Val Leu Thr Gly Ala 50 55 60	434
GGC GAA TCC TGG ACC GCG GGC ATG GAC CTG AAG GAG TAT TTC CGC GAG Gly Glu Ser Trp Thr Ala Gly Met Asp Leu Lys Glu Tyr Phe Arg Glu 65 70 75	482
ACC GAT GCT GGC CCC GAA ATT CTG CAA GAG AAG ATT CGT CGGGGACAGC Thr Asp Ala Gly Pro Glu Ile Leu Gln Glu Lys Ile Arg 80 85 90 91	531
AAGCGAACCG GAATTGCCAG CTGGGGCGCC CTCTGGTAAG GTTGGGAAGC CCTGCAAAGT	591
AAACTGGATG GCTTTCTTGC CGCCAAGGAT CTGATGGCGC AGGGGATCAA GATCTGATCA	651
AGAGACAGGA TGAGGATCGT TTCGC ATG ATT GAA CAA GAT GGA TTG CAC GCA Met Ile Glu Gln Asp Gly Leu His Ala 1 5	703
GGT TCT CCG GCC GCT TGG GTG GAG AGG CTA TTC GGC TAT GAC TGG GCA Gly Ser Pro Ala Ala Trp Val Glu Arg Leu Phe Gly Tyr Asp Trp Ala 10 20 25	751
CAA CAG ACA ATC GGC TGC TCT GAT GCC GCC GTG TTC CGG CTG TCA GCG Gln Gln Thr Ile Gly Cys Ser Asp Ala Ala Val Phe Arg Leu Ser Ala 30 35 40	799
CAG GGG CGC CCG GTT CTT TTT GTC AAG ACC GAC CTG TCC GGT GCC CTG Gln Gly Arg Pro Val Leu Phe Val Lys Thr Asp Leu Ser Gly Ala Leu 45	847
AAT GAA CTG CAG GAC GAG GCA GCG CGG CTA TCG TGG CTG GCC ACG ACG ASn Glu Leu Gln Asp Glu Ala Ala Arg Leu Ser Trp Leu Ala Thr Thr 60 65 70	895

		?
		•
		•

	-				GCT Ala										943
					GGC Gly 95										991
					GAG Glu										1039
					GAT Asp										1087
					CGA Arg										1135
					GAC Asp										1183
		-			AAG Lys 175										1231
					GCC Ala										1279
					ATC Ile										1327
			Ile		TTG Leu								 		1375
		Trp			CGC Arg							Ile			1423
	Ser				GCC Ala 255	Phe					Asp			TGA	1471
GCG	GGAC	TCT	GGGG	TTCG	AA A	TGAC	CGAC	C AA	.GCGA	.CGCC		GAG Glu 255			1525
		Phe					Ser					Leu		TAC Tyr	1573

1
•
i i
·
<b>‹</b>

AAG CGC TGA Lys Arg 275 276	TAAATGCGCC	: GGGGCCCTCG	CTGCGCCCCC	GGCCTTCCAA	TAATGACAAT	1632
AATGAGGAGT	GCCCAATGTT	TCACGTGCCC	CTGCTTATTG	GTGGTAAGCC	TTGTTCAGCA	1692
TCTGATGAGC	GCACCTTCGA	GCGTCGTAGC	CCGCTGACCG	GAGAAGTGGT	ATCGCGCGTC	1752
GCTGCTGCCA	GTTTGGAAGA	TGCGGACGCC	GCAGTGGCCG	CTGCACAGGC	TGCGTTTCCT	1,812
GAATGGGCGG	CGCTTGCTCC	GAGCGAACGC	CGTGCCCGAC	TGCTGCGAGC	GGCGGATCTT	1872
CTAGAGGACC	GTTCTTCCGA	GTTCACCGCC	GCAGCGAGTG	AAACTGGCGC	AGCGGGAAAC	1932
TGGTATGGGT	TTAACGTTTA	CCTGGCGGCG	GGCATGTTGC	GGGGAATTC		1981
FIG. 2j:						

	4
	•
	-
	•
4	
	4
	4
	•

GAATTCCCCT GGCGACGAAA GGGCGGCAGG CCGCATGGCC ACGGCTGGGC GGTAACTGAT	60
	00
GCTTGCGTTA ATCGTTAACC GTTTGAAATT CCTTGCCAAA TTTCGGCGAG AGAATCATGC	120
GGGTACGCCT TTCCGTGCGC TTTGATCTGC GCTTCCGTGC CTTGAATCAG AAAAATAGTT	180
AATTGACAGA ACTATAGGTT CGCAGTAGCT TTTGCTCACC CACCAAATCC ACAGCACTGG	240
GGTGCACG ATG AAT AGC TAC GAT GGC CGT TGG TCT ACC GTT GAT GTG AAG  Met Asn Ser Tyr Asp Gly Arg Trp Ser Thr Val Asp Val Lys  1 5 10	290
GTT GAA GAA GGT ATC GCT TGG GTC ACG CTG AAC CGC CCG GAG AAG CGC Val Glu Glu Gly Ile Ala Trp Val Thr Leu Asn Arg Pro Glu Lys Arg 15 20 25 30	338
AAC GCA ATG AGC CCA ACT CTC AAT CGA GAG ATG GTC GAG GTT CTG GAG Asn Ala Met Ser Pro Thr Leu Asn Arg Glu Met Val Glu Val Leu Glu 35 40 45	386
GTG CTG GAG CAG GAC GCA GAT GCT CGC GTG CTT GTT CTG ACT GGT GCA Val Leu Glu Gln Asp Ala Asp Ala Arg Val Leu Val Leu Thr Gly Ala 50 55 60	434
GGC GAA TCC TGG ACC GCG GGC ATG GAC CTG AAG GAG TAT TTC CGC GAG Gly Glu Ser Trp Thr Ala Gly Met Asp Leu Lys Glu Tyr Phe Arg Glu 65	482
ACC GAT GCT GGC CCC GAA ATT CTG CAA GAG AAG ATT CGT CGGGGGAGAG Thr Asp Ala Gly Pro Glu Ile Leu Gln Glu Lys Ile Arg 80 85 90 91	531
GCGGTTTGCG TATTGGGCGC ATGCATAAAA ACTGTTGTAA TTCATTAAGC ATTCTGCCGA	591
CATGGAAGCC ATCACAAACG GCATGATGAA CCTGAATCGC CAGCGGCATC AGCACCTTGT	651
CGCCTTGCGT ATAATATTTG CCCATGGACG CACACCGTGG AAACGGATGA AGGCACGAAC	711
CCAGTTGACA TAAGCCTGTT CGGTTCGTAA ACTGTAATGC AAGTAGCGTA TGCGCTCACG	771
CAACTGGTCC AGAACCTTGA CCGAACGCAG CGGTGGTAAC GGCGCAGTGG CGGTTTTCAT	831
GGCTTGTTAT GACTGTTTTT TTGTACAGTC TATGCCTCGG GCATCCAAGC AGCAAGCGCG	891
TTACGCCGTG GGTCGATGTT TGATGTTATG GAGCAGCAAC G ATG TTA CGC AGC AGC Met Leu Arg Ser Ser 1	947
AAC GAT GTT ACG CAG CAG GGC AGT CGC CCT AAA ACA AAG TTA GGT GGC Asn Asp Val Thr Gln Gln Gly Ser Arg Pro Lys Thr Lys Leu Gly Gly 10 15 20	995
TCA AGT ATG GGC ATC ATT CGC ACA TGT AGG CTC GGC CCT GAC CAA GTC Ser Ser Met Gly Ile Ile Arg Thr Cys Arg Leu Gly Pro Asp Gln Val 25 30 35	1043

		•
		•
		1
		•

AAA TCC ATG CGG GCT GCT CTT GAT CTT TTC GGT CGT GAG TTC GGA GAC Lys Ser Met Arg Ala Ala Leu Asp Leu Phe Gly Arg Glu Phe Gly Asp 40 45 50	1091
GTA GCC ACC TAC TCC CAA CAT CAG CCG GAC TCC GAT TAC CTC GGG AAC Val Ala Thr Tyr Ser Gln His Gln Pro Asp Ser Asp Tyr Leu Gly Asn 55 60 65	1139
TTG CTC CGT AGT AAG ACA TTC ATC GCG CTT GCT GCC TTC GAC CAA GAA Leu Leu Arg Ser Lys Thr Phe Ile Ala Leu Ala Ala Phe Asp Gln Glu 70 75 80 85	1187
GCG GTT GTT GGC GCT CTC GCG GCT TAC GTT CTG CCC AGG TTT GAG CAG Ala Val Val Gly Ala Leu Ala Ala Tyr Val Leu Pro Arg Phe Glu Gln 90 95 100	1235
CCG CGT AGT GAG ATC TAT ATC TAT GAT CTC GCA GTC TCC GGC GAG CAC Pro Arg Ser Glu Ile Tyr Ile Tyr Asp Leu Ala Val Ser Gly Glu His 105	1283
CGG AGG CAG GGC ATT GCC ACC GCG CTC ATC AAT CTC CTC AAG CAT GAG Arg Arg Gln Gly Ile Ala Thr Ala Leu Ile Asn Leu Leu Lys His Glu 120 125 130	1331
GCC AAC GCG CTT GGT GCT TAT GTG ATC TAC GTG CAA GCA GAT TAC GGT Ala Asn Ala Leu Gly Ala Tyr Val Ile Tyr Val Gln Ala Asp Tyr Gly 135	1379
GAC GAT CCC GCA GTG GCT CTC TAT ACA AAG TTG GGC ATA CGG GAA GAA Asp Asp Pro Ala Val Ala Leu Tyr Thr Lys Leu Gly Ile Arg Glu Glu 150 160 165	1427
GTG ATG CAC TTT GAT ATC GAC CCA AGT ACC GCC ACC TAA CAATTCGTTC Val Met His Phe Asp Ile Asp Pro Ser Thr Ala Thr 170 175 177	1476
AAGCCGAGAT CGGCTTCCCC GAG CAG GGC ATG AAG CAG TTC CTT GAC GAG Glu Gln Gly Met Lys Gln Phe Leu Asp Glu 255 260	1526
AAA AGC ATC AAG CCG GGC TTG CAG ACC TAC AAG CGC TGA TAAATGCGCC Lys Ser Ile Lys Pro Gly Leu Gln Thr Tyr Lys Arg 265 270 275 276	1575
GGGGCCCTCG CTGCGCCCCC GGCCTTCCAA TAATGACAAT AATGAGGAGT GCCCAATGTT	1635
TCACGTGCCC CTGCTTATTG GTGGTAAGCC TTGTTCAGCA TCTGATGAGC GCACCTTCGA	1695
GCGTCGTAGC CCGCTGACCG GAGAAGTGGT ATCGCGCGTC GCTGCTGCCA GTTTGGAAGA	1755
TGCGGACGCC GCAGTGGCCG CTGCACAGGC TGCGTTTCCT GAATGGGCGG CGCTTGCTCC	1815
GAGCGAACGC CGTGCCCGAC TGCTGCGAGC GGCGGATCTT CTAGAGGACC GTTCTTCCGA	1875

		•

GTTCACCGCC GCAGCGAGTG AAACTGGCGC AGCGGGAAAC TGGTATGGGT TTAACGTTTA 1935

CCTGGCGGCG GGCATGTTGC GGGGAATTC 1964

FIG. 2k:

29

PCT/EP99/07952

WO 00/26355

		•
		•
		•

GAATTCCCCT GGCGACGAAA GGGCGGCAGG CCGCATGGCC ACGGCTGGGC GGTAACTGAT	60
GCTTGCGTTA ATCGTTAACC GTTTGAAATT CCTTGCCAAA TTTCGGCGAG AGAATCATGC	120
GGGTACGCCT TTCCGTGCGC TTTGATCTGC GCTTCCGTGC CTTGAATCAG AAAAATAGTT	180
AATTGACAGA ACTATAGGTT CGCAGTAGCT TTTGCTCACC CACCAAATCC ACAGCACTGG	240
GGTGCACG ATG AAT AGC TAC GAT GGC CGT TGG TCT ACC GTT GAT GTG AAG  Met Asn Ser Tyr Asp Gly Arg Trp Ser Thr Val Asp Val Lys  1 5 10	290
GTT GAA GAA GGT ATC GCT TGG GTC ACG CTG AAC CGC CCG GAG AAG CGC Val Glu Glu Gly Ile Ala Trp Val Thr Leu Asn Arg Pro Glu Lys Arg 25 30	338
AAC GCA ATG AGC CCA ACT CTC AAT CGA GAG ATG GTC GAG GTT CTG GAG Asn Ala Met Ser Pro Thr Leu Asn Arg Glu Met Val Glu Val Leu Glu 35 40 45	386
GTG CTG GAG CAG GAC GCA GAT GCT CGC GTG CTT GTT CTG ACT GGT GCA Val Leu Glu Gln Asp Ala Asp Ala Arg Val Leu Val Leu Thr Gly Ala 50 55 60	434
GGC GAA TCC TGG ACC GCG GGC ATG GAC CTG AAG GAG TAT TTC CGC GAG Gly Glu Ser Trp Thr Ala Gly Met Asp Leu Lys Glu Tyr Phe Arg Glu 65 70 75	482
ACC GAT GCT GGC CCC GAA ATT CTG CAA GAG AAG ATT CGT CGC GAG CAG Thr Asp Ala Gly Pro Glu Ile Leu Gln Glu Lys Ile Arg Arg Glu Gln 80 85 90 92 255	530
GGC ATG AAG CAG TTC CTT GAC GAG AAA AGC ATC AAG CCG GGC TTG CAG Gly Met Lys Gln Phe Leu Asp Glu Lys Ser Ile Lys Pro Gly Leu Gln 260 265 270	578
ACC TAC AAG CGC TGA TAAATGCGCC GGGGCCCTCG CTGCGCCCCC GGCCTTCCAA Thr Tyr Lys Arg 275 276	633
TAATGACAAT AATGAGGAGT GCCCAATGTT TCACGTGCCC CTGCTTATTG GTGGTAAGCC	693
TTGTTCAGCA TCTGATGAGC GCACCTTCGA GCGTCGTAGC CCGCTGACCG GAGAAGTGGT	753
ATCGCGCGTC GCTGCTGCCA GTTTGGAAGA TGCGGACGCC GCAGTGGCCG CTGCACAGGC	813
TGCGTTTCCT GAATGGGCGG CGCTTGCTCC GAGCGAACGC CGTGCCCGAC TGCTGCGAGC	873
GGCGGATCTT CTAGAGGACC GTTCTTCCGA GTTCACCGCC GCAGCGAGTG AAACTGGCGC	933
AGCGGGAAAC TGGTATGGGT TTAACGTTTA CCTGGCGGCG GGCATGTTGC GGGGAATTC	992
FIG. 21:	

•
•
•

GAATT	'CCA	AT A	ATGA	CAAT	А АТ	GAGG	AGTG	ccc	Ме				l Pr		G CTT u Leu	55
ATT G	GT (	GGT Gly 10	AAG Lys	CCT Pro	TGT Cys	TCA Ser	GCA Ala 15	TCT Ser	GAT Asp	GAG Glu	CGC Arg	ACC Thr 20	TTC Phe	GAG Glu	CGT Arg	103
CGT A	AGC Ser 25	CCG Pro	CTG Leu	ACC Thr	GGA Gly	GAA Glu 30	GTG Val	GTA Val	TCG Ser	CGC Arg	GTC Val 35	GCT Ala	GCT Ala	GCC Ala	AGT Ser	151
TTG G Leu G 40	GAA Glu	GAT Asp	GCG Ala	GAC Asp	GCC Ala 45	GCA Ala	GTG Val	GCC Ala	GCT Ala	GCA Ala 50	CAG Gln	GCT Ala	GCG Ala	TTT Phe	CCT Pro 55	199
GAA 1 Glu 1	rgg Irp	GCG Ala	GCG Ala	CTT Leu 60	GCT Ala	CCG Pro	AGC Ser	GAA Glu	CGC Arg 65	CGT Arg	GCC Ala	CGA Arg	CTG Leu	CTG Leu 70	CGA Arg	247
GCG (	GCG Ala	GAT Asp	CTT Leu 75	CTA Leu	GAG Glu	GAC Asp	CGT Arg	TCT Ser 80	TCC Ser	GAG Glu	TTC Phe	ACC Thr	GCC Ala 85	GCA Ala	GCG Ala	295
AGT (	GAA Glu	ACT Thr 90	GGC Gly	GCA Ala	GCG Ala	GGA Gly	AAC Asn 95	TGG Trp	TAT Tyr	GGG Gly	TTT Phe	AAC Asn 100	GTT Val	TAC Tyr	CTG Leu	343
GCG (	GCG Ala 105	GGC Gly	ATG Met	TTG Leu	CGG Arg	GAA Glu 110	GCC Ala	GCG Ala	GCC Ala	ATG Met	ACC Thr 115	ACA Thr	CAG Gln	ATT Ile	CAG Gln	391
GGC Gly 120	GAT Asp	GTC Val	ATT Ile	CCG Pro	TCC Ser 125	AAT Asn	GTG Val	CCC Pro	GGT Gly	AGC Ser 130	TTT Phe	GCC Ala	ATG Met	GCG Ala	GTT Val 135	439
CGA Arg	CAG Gln	CCA Pro	TGT Cys	GGC Gly 140	Val	GTG Val	CTC Leu	GGT Gly	ATT Ile 145	Ala	CCT Pro	TGG Trp	AAT Asn	GCT Ala 150	CCG Pro	487
GTA Val	ATC Ile	CTT Leu	GGC Gly 155	Val	. CGG Arg	GCT Ala	GTT Val	GCG Ala 160	Met	CCG Pro	TTG Leu	GCA Ala	TGC Cys 165	GGC Gly	AAT Asn	535
ACC Thr	GTG Val	GTG Val 170	Leu	AAA Lys	AGC Ser	TCT Ser	GAG Glu 175	Leu	AGT Ser	CCC Pro	TTT Phe	ACC Thr 180	His	CGC Arg	CTG Leu	583
ATT Ile	GGT Gly 185	Glr	GTG Val	TTG Leu	CAT His	GAT Asp 190	Ala	Cl)	CTG Leu	GGG Gly	GAT Asp 195	Gly	GTG Val	GTG Val	AAT Asn	631
GTC Val 200	ATC Ile	AG0 Sei	AAT Asr	GCC Ala	205	Glr	A GAC n Asp	GCT Ala	CCT a Pro	GCG Ala 210	val	GTG Val	GAG Glu	CGA Arc	CTG Leu 215	679

2
•
•
1 <u>2</u>
•
ů.
i

ATT Ile	GCA Ala	AAT Asn	CCT Pro	GCG Ala 220	GTA Val	CGT Arg	CGA Arg	GTG Val	AAC Asn 225	TTC Phe	ACC Thr	GGT Gly	TCG Ser	ACC Thr 230	CAC His	727
GTT Val	GGA Gly	CGG Arg	ATC Ile 235	ATT Ile	GGT Gly	GAG Glu	CTG Leu	TCT Ser 240	GCG Ala	CGT Arg	CAT His	CTG Leu	AAG Lys 245	CCT Pro	GCT Ala	775
GTG Val	CTG Leu	GAA Glu 250	TTA Leu	GGT Gly	GGT Gly	AAG Lys	GCT Ala 255	CCG Pro	TTC Phe	TTG Leu	GTC Val	TTG Leu 260	GAC Asp	GAT Asp	GCC Ala	823
GAC Asp	CTC Leu 265	GAT Asp	GCG Ala	GCG Ala	GTC Val	GAA Glu 270	GCG Ala	GCG Ala	GCC Ala	TTT Phe	GGT Gly 275	GCC Ala	TAC Tyr	TTC Phe	AAT Asn	871
CAG Gln 280	GGT Gly	CAA Gln	ATC Ile	TGC Cys	ATG Met 285	TCC Ser	ACT Thr	GAG Glu	CGT Arg	CTG Leu 290	ATT Ile	GTG Val	ACA Thr	GCA Ala	GTC Val 295	919
GCA Ala	GAC Asp	GCC Ala	TTT Phe	GTT Val 300	Glu	AAG Lys	CTG Leu	GCG Ala	AGG Arg 305	AAG Lys	GTC Val	GCC Ala	ACA Thr	CTG Leu 310	CGT Arg	967
GCT Ala	GGC Gly	GAT Asp	CCT Pro	Asn	GAT Asp	CCG Pro	CAA Gln	TCG Ser 320	Val	TTG Leu	GGT Gly	TCG Ser	TTG Leu 325	Ile	GAT Asp	1015
GCC Ala	AA1 Asr	GCA Ala 330	Gly	CAA Gln	CGC Arg	ATC Ile	CAG Glr 335	Val	CTG	GTC Val	GAT Asp	GAT Asp 340	Ala	CTC Leu 342	GGG	1063
GAC	AGC	AGC	GAAC	CGGP	T TAL	GCCA	GCT	G GG	GCGCC	CTCI	GGT	AAGO	STTG	GGAA	GCCCTG	1123
CAA	AGT <i>I</i>	AAAC	TGG	ATGGC	TT T	CTTG	CCGC	CC AF	AGGAT	CTG	TGC	GCGC <i>I</i>	AGGG	GATO	CAAGATC	1183
TGA	TCA	AGAG	ACAG	GGATO	GAG C	ATCO	STTT(	CG C	ATG Met 1	ATT Ile	GAA Glu	CAA Gln	GAT Asp 5	GGA Gly	TTG Leu	1235
CAC His	GC Ala	A GG a Gl	y Se	r CCC	G GCC	C GCT	TGG Trj	o Va	G GA0 l Glu	AGO Aro	G CTA	A TTO D Pho 2	e Gl	С ТАТ у Туз	r GAC r Asp	1283
TG( Trp	G GC Al 2	a Gl	A CA n Gl	G AC	A ATO	GGG Gl <sub>2</sub> 3	у Су	C TC s Se	T GA' r As <sub>l</sub>	r GCO p Al	C GCG a Al	a Va	G TT	C CGG	G CTG g Leu	1331
TCA Sea	r Al	G CA a Gl	G GG n Gl	G CG y Ar	C CCC g Pr	o Va	T CT l Le	T TT u Ph	T GT e Va	C AA 1 Ly 5	s Th	C GA r As	C CT p Le	G TC u Se	C GGT r Gly 55	1379
GC(	C CT a Le	G AA u As	T GA n Gl	u Le	G CA u Gl 0	G GA n As	C GA p Gl	G GC u Al	a Al	G CG a Ar 5	G CT g Le	A TC u Se	G TG	p Le	G GCC u Ala O	1427

•
•
•

33

ACG Thr																1475
GGA Gly	AGG Arg	GAC Asp 90	TGG Trp	CTG Leu	CTA Leu	TTG Leu	GGC Gly 95	GAA Glu	GTG Val	CCG Pro	GGG Gly	CAG Gln 100	GAT Asp	CTC Leu	CTG Leu	1523
			CTT Leu													1571
ATG Met 120	CGG Arg	CGG Arg	CTG Leu	CAT His	ACG Thr 125	CTT Leu	GAT Asp	CCG Pro	GCT Ala	ACC Thr 130	TGC Cys	CCA Pro	TTC Phe	GAC Asp	CAC His 135	1619
CAA Gln	GCG Ala	AAA Lys	CAT His	CGC Arg 140	ATC Ile	GAG Glu	CGA Arg	GCA Ala	CGT Arg 145	ACT Thr	CGG Arg	ATG Met	GAA Glu	GCC Ala 150	GGT Gly	1667
CTT Leu	GTC Val	GAT Asp	CAG Gln 155	GAT Asp	GAT Asp	CTG Leu	GAC Asp	GAA Glu 160	GAG Glu	CAT	CAG Gln	GGG Gly	CTC Leu 165	GCG Ala	CCA Pro	1715
Ala	Glu	Leu 170	TTC Phe	Ala	Arg	Leu	Lys 175	Ala	Arg	Met	Pro	Asp 180	Gly	Glu	Asp	1763
Leu	Val 185	Val	Thr	His	Gly	Asp 190	Ala	Cys	Leu	Pro	Asn 195	Ile	Met	Val		1811
Asn 200	Gly	Arg	Phe	Ser	Gly 205	Phe	Ile	Asp	Cys	Gly 210	Arg	Leu	Gly	Val	GCG Ala 215	1859
Asp	Arg	Tyr	Gln	220	lle	Ala	Leu	Ala	Thr 225	Arg	, Asp	Ile	Ala	Glu 230		1907
Leu	Gly	, Gl	/ Glu 235	Trp	Ala	Asp	Arg	Phe 240	e Leu	ı Val	L <b>L</b> ev	туг	Gly 245	, Il€	GCC Ala	1955
Ala	Pro	250	Ser	Glr	n Arg	, Ile	255	a Ph∈	э Туг	Arq	g Lei	260	a Asp	o Glu	TTC Phe	2003
Phe 264	) 1		GGAC											Ala 421	Gln	2057
CG( Ar	C GTO g Val	C GA' l As <sub>i</sub> 42	p Se	G GGG	C AT	r TG0 e Cy:	C CA? 5 His 430	s Ile	C AA' e Ası	r GG n Gl	A CC	G AC' o Th: 43	r Vai	G CA' L Hi:	r GAC s Asp	2105

•
•
***
•

Glu	GCT Ala 440	CAG Gln	ATG Met	CCA Pro	TTC Phe	GGT Gly 445	GGG Gly	GTG Val	AAG Lys	TCC Ser	AGC Ser 450	GGC Gly	TAC Tyr	GGC Gly	AGC Ser	2153
TTC Phe 455	GGC Gly	AGT Ser	CGA Arg	GCA Ala	TCG Ser 460	ATT Ile	GAG Glu	CAC His	TTT Phe	ACC Thr 465	CAG Gln	CTG Leu	CGC Arg	TGG Trp	CTG Leu 470	2201
				GGC Gly 475					Pro		TAA	ATC	GATC'	TTC		2247
GGGG	CGCC	GCG	GGCA	TCAT	GC C	CGCG	GCGC'	r CG	CCTC	TTTA	CAA	TCTC'	TAA	CTTG	AAAATA	2307
ACAC	GAGC'	TGT	TCTC	CGGT	CT T	GGTG	GATC	A AG	GCCA	STCG	CGG.	AGAG'	TCT	CGAA	GAGGAG	2367
AGT	ACAG'	TGA	ACGC	CGAG	TC C	ACAT	TGCA	A CC	GCAG	GCAT	CAT	CATG	CTC	TGCT	CAGCCA	2427
CGC	racc	GCA	GTGT	GTCG	ат т	GGTC	ATCC	T CC	GGTT	GAGG	TTA	CGCA	AGA	CGCT	GGAGGT	2487
ATT	GTCC	GGA	TGCG	TTCT	CT C	GAGG	CGCT	т ст	TCCC	TTCC	CGG	GTGG	AAT	TC		2539
FIG	. 2m	:														

•
•
•
•

GAATTCCAAT AATGACAATA ATGAGGAGTG CCCA ATG TTT CAC GTG CCC CTG CTT 55 Met Phe His Val Pro Leu Leu 7 ATT GGT GGT AAG CCT TGT TCA GCA TCT GAT GAG CGC ACC TTC GAG CGT 103 Ile Gly Gly Lys Pro Cys Ser Ala Ser Asp Glu Arg Thr Phe Glu Arg 15 10 CGT AGC CCG CTG ACC GGA GAA GTG GTA TCG CGC GTC GCT GCC AGT 151 Arg Ser Pro Leu Thr Gly Glu Val Val Ser Arg Val Ala Ala Ala Ser 25 TTG GAA GAT GCG GAC GCC GCA GTG GCC GCT GCA CAG GCT GCG TTT CCT 199 Leu Glu Asp Ala Asp Ala Ala Val Ala Ala Ala Gln Ala Ala Phe Pro 50 45 GAA TGG GCG GCG CTT GCT CCG AGC GAA CGC CGT GCC CGA CTG CTG CGA 247 Glu Trp Ala Ala Leu Ala Pro Ser Glu Arg Arg Ala Arg Leu Leu Arg 60 GCG GCG GAT CTT CTA GAG GAC CGT TCT TCC GAG TTC ACC GCC GCA GCG 295 Ala Ala Asp Leu Leu Glu Asp Arg Ser Ser Glu Phe Thr Ala Ala Ala AGT GAA ACT GGC GCA GCG GGA AAC TGG TAT GGG TTT AAC GTT TAC CTG 343 Ser Glu Thr Gly Ala Ala Gly Asn Trp Tyr Gly Phe Asn Val Tyr Leu 90 GCG GCG GGC ATG TTG CGG GAA GCC GCG GCC ATG ACC ACA CAG ATT CAG 391 Ala Ala Gly Met Leu Arg Glu Ala Ala Ala Met Thr Thr Gln Ile Gln 110 105 GGC GAT GTC ATT CCG TCC AAT GTG CCC GGT AGC TTT GCC ATG GCG GTT 439 Gly Asp Val Ile Pro Ser Asn Val Pro Gly Ser Phe Ala Met Ala Val 130 125 120 CGA CAG CCA TGT GGC GTG GTG CTC GGT ATT GCG CCT TGG AAT GCT CCG 487 Arg Gln Pro Cys Gly Val Val Leu Gly Ile Ala Pro Trp Asn Ala Pro 140 GTA ATC CTT GGC GTA CGG GCT GTT GCG ATG CCG TTG GCA TGC GGC AAT 535 Val Ile Leu Gly Val Arg Ala Val Ala Met Pro Leu Ala Cys Gly Asn 155 ACC GTG GTG TTG AAA AGC TCT GAG CTG AGT CCC TTT ACC CAT CGC CTG 583 Thr Val Val Leu Lys Ser Ser Glu Leu Ser Pro Phe Thr His Arg Leu 170 ATT GGT CAG GTG TTG CAT GAT GCT GGT CTG GGG GAT GGC GTG GTG AAT 631 Ile Gly Gln Val Leu His Asp Ala Gly Leu Gly Asp Gly Val Val Asn 190 185 GTC ATC AGC AAT GCC CCG CAA GAC GCT CCT GCG GTG GTG GAG CGA CTG 679 Val Ile Ser Asn Ala Pro Gln Asp Ala Pro Ala Val Val Glu Arg Leu 205 200

	•
	•
	•
	٠

ATT GCA AAT CCT GCG GTA CGT CGA GTG AAC TTC ACC GGT TCG ACC CF Ile Ala Asn Pro Ala Val Arg Arg Val Asn Phe Thr Gly Ser Thr Hi 220 225 230	AC 727 is
GTT GGA CGG ATC ATT GGT GAG CTG TCT GCG CGT CAT CTG AAG CCT GC Val Gly Arg Ile Ile Gly Glu Leu Ser Ala Arg His Leu Lys Pro Al 235	CT 775 la
GTG CTG GAA TTA GGT GGT AAG GCT CCG TTC TTG GTC TTG GAC GAT GC Val Leu Glu Leu Gly Gly Lys Ala Pro Phe Leu Val Leu Asp Asp Asp 250	CC 823 la
GAC CTC GAT GCG GCG GTC GAA GCG GCG GCC TTT GGT GCC TAC TTC AA Asp Leu Asp Ala Ala Val Glu Ala Ala Ala Phe Gly Ala Tyr Phe As 265 270 275	AT 871 sn
CAG GGT CAA ATC TGC ATG TCC ACT GAG CGT CTG ATT GTG ACA GCA GC GIn Gly Gln Ile Cys Met Ser Thr Glu Arg Leu Ile Val Thr Ala Van 280 285 290 290	TC 919 al 95
GCA GAC GCC TTT GTT GAA AAG CTG GCG AGG AAG GTC GCC ACA CTG CC Ala Asp Ala Phe Val Glu Lys Leu Ala Arg Lys Val Ala Thr Leu A 300 305 310	GT 967 .rg
GCT GGC GAT CCT AAT GAT CCG CAA TCG GTC TTG GGT TCG TTG ATT GAT GAT GAT GAT GAT GAT GAT GA	AT 1015 sp
GCC AAT GCA GGT CAA CGC ATC CAG GTGGGGAGAG GCGGTTTGCG TATTGGG Ala Asn Ala Gly Gln Arg Ile Gln 330 335	SCGC 1069
ATGCATAAAA ACTGTTGTAA TTCATTAAGC ATTCTGCCGA CATGGAAGCC ATCACA	AAACG 1129
GCATGATGAA CCTGAATCGC CAGCGGCATC AGCACCTTGT CGCCTTGCGT ATAATA	ATTTG 1189
CCCATGGACG CACACCGTGG AAACGGATGA AGGCACGAAC CCAGTTGACA TAAGCC	CTGTT 1249
CGGTTCGTAA ACTGTAATGC AAGTAGCGTA TGCGCTCACG CAACTGGTCC AGAACC	CTTGA 1309
CCGAACGCAG CGGTGGTAAC GGCGCAGTGG CGGTTTTCAT GGCTTGTTAT GACTGT	TTTTT 1369
TTGTACAGTC TATGCCTCGG GCATCCAAGC AGCAAGCGCG TTACGCCGTG GGTCGA	ATGTT 1429
TGATGTTATG GAGCAGCAAC G ATG TTA CGC AGC AGC AAC GAT GTT ACG (  Met Leu Arg Ser Asn Asp Val Thr (  1 5	CAG 1480 Gln 10
CAG GGC AGT CGC CCT AAA ACA AAG TTA GGT GGC TCA AGT ATG GGC AGT Gln Gly Ser Arg Pro Lys Thr Lys Leu Gly Gly Ser Ser Met Gly 15	ATC 1528 Ile
ATT CGC ACA TGT AGG CTC GGC CCT GAC CAA GTC AAA TCC ATG CGG ( Ile Arg Thr Cys Arg Leu Gly Pro Asp Gln Val Lys Ser Met Arg 3 30 35 40	GCT 1576 Ala

•
2
-
•

GCT CTT GAT CTT TTC GGT CGT GAG TTC GGA GAC GTA GCC ACC TAC TCC	1624
Ala Leu Asp Leu Phe Gly Arg Glu Phe Gly Asp Val Ala Thr Tyr Ser 45 50 55	
CAA CAT CAG CCG GAC TCC GAT TAC CTC GGG AAC TTG CTC CGT AGT AAG Gln His Gln Pro Asp Ser Asp Tyr Leu Gly Asn Leu Leu Arg Ser Lys 60 65 70	1672
ACA TTC ATC GCG CTT GCT GCC TTC GAC CAA GAA GCG GTT GTT GGC GCT Thr Phe Ile Ala Leu Ala Ala Phe Asp Gln Glu Ala Val Val Gly Ala 75 80 85 90	1,720
CTC GCG GCT TAC GTT CTG CCC AGG TTT GAG CAG CCG CGT AGT GAG ATC Leu Ala Ala Tyr Val Leu Pro Arg Phe Glu Gln Pro Arg Ser Glu Ile 95 100 105	1768
TAT ATC TAT GAT CTC GCA GTC TCC GGC GAG CAC CGG AGG CAG GGC ATT  Tyr Ile Tyr Asp Leu Ala Val Ser Gly Glu His Arg Arg Gln Gly Ile  110 115 120	1816
GCC ACC GCG CTC ATC AAT CTC CTC AAG CAT GAG GCC AAC GCG CTT GGT Ala Thr Ala Leu Ile Asn Leu Leu Lys His Glu Ala Asn Ala Leu Gly 125 130 135	1864
GCT TAT GTG ATC TAC GTG CAA GCA GAT TAC GGT GAC GAT CCC GCA GTG Ala Tyr Val Ile Tyr Val Gln Ala Asp Tyr Gly Asp Asp Pro Ala Val 140 145 150	1912
GCT CTC TAT ACA AAG TTG GGC ATA CGG GAA GAA GTG ATG CAC TTT GAT Ala Leu Tyr Thr Lys Leu Gly Ile Arg Glu Glu Val Met His Phe Asp 155 160 165 170	1960
ATC GAC CCA AGT ACC GCC ACC TAA CAATTCGTTC AAGCCGAGAT CGGCTTCCCA Ile Asp Pro Ser Thr Ala Thr 175 177	2014
A TTG GCC CAG CGC GTC GAT TCG GGC ATT TGC CAT ATC AAT GGA CCG ACT Leu Ala Gln Arg Val Asp Ser Gly Ile Cys His Ile Asn Gly Pro Thr 420 425 430 435	2063
GTG CAT GAC GAG GCT CAG ATG CCA TTC GGT GGG GTG AAG TCC AGC GGC Val His Asp Glu Ala Gln Met Pro Phe Gly Gly Val Lys Ser Ser Gly 440 445 450	2111
TAC GGC AGC TTC GGC AGT CGA GCA TCG ATT GAG CAC TTT ACC CAG CTG Tyr Gly Ser Phe Gly Ser Arg Ala Ser Ile Glu His Phe Thr Gln Leu 455 460 465	2159
CGC TGG CTG ACC ATT CAG AAT GGC CCG CGG CAC TAT CCA ATC TAA Arg Trp Leu Thr Ile Gln Asn Gly Pro Arg His Tyr Pro Ile 470 475 480 481	2204
ATCGATCTTC GGGCGCCGCG GGCATCATGC CCGCGGCGCT CGCCTCATTT CAATCTCTAA	2264
CTTGATAAAA ACAGAGCTGT TCTCCGGTCT TGGTGGATCA AGGCCAGTCG CGGAGAGTCT	2324

•
•
•
•

WO 00/26355 PCT/EP99/07952

CGAAGAGGAG	AGTACAGTGA	ACGCCGAGTC	CACATTGCAA	CCGCAGGCAT	CATCATGCTC	2384
TGCTCAGCCA	CGCTACCGCA	GTGTGTCGAT	TGGTCATCCT	CCGGTTGAGG	TTACGCAAGA	2444
CGCTGGAGGT	ATTGTCCGGA	TGCGTTCTCT	CGAGGCGCTT	CTTCCCTTCC	CGGGTGGAAT	2504
TC						2506
FIG. 2n:						

				,
				•
				į

GAATT	CCA	AT A	ATGA	CAAT	A AT	GAGG.	AGTG	ccc	A AT Me	G TT t Ph 1	T CA e Hi	C GT S Va	l Pr	C CT o Le 5	G CTT u Leu	55
ATT G	GT Gly	GGT Gly 10	AAG Lys	CCT Pro	TGT Cys	TCA Ser	GCA Ala 15	TCT Ser	GAT Asp	GAG Glu	CGC Arg	ACC Thr 20	TTC Phe	GAG Glu	CGT Arg	103
CGT A	AGC Ser 25	CCG Pro	CTG Leu	ACC Thr	GGA Gly	GAA Glu 30	GTG Val	GTA Val	TCG Ser	CGC Arg	GTC Val 35	GCT Ala	GCT Ala	GCC Ala	AGT Ser	151
TTG ( Leu ( 40	GAA Glu	GAT Asp	GCG Ala	GAC Asp	GCC Ala 45	GCA Ala	GTG Val	GCC Ala	GCT Ala	GCA Ala 50	CAG Gln	GCT Ala	GCG Ala	TTT Phe	CCT Pro 55	199
GAA 1	rgg Trp	GCG Ala	GCG Ala	CTT Leu 60	GCT Ala	CCG Pro	AGC Ser	GAA Glu	CGC Arg 65	CGT Arg	GCC Ala	CGA Arg	CTG Leu	CTG Leu 70	CGA Arg	247
GCG (	GCG Ala	GAT Asp	CTT Leu 75	CTA Leu	GAG Glu	GAC Asp	CGT Arg	TCT Ser 80	TCC Ser	GAG Glu	TTC Phe	ACC Thr	GCC Ala 85	GCA Ala	GCG Ala	295
AGT (	GAA Glu	ACT Thr 90	GGC Gly	GCA Ala	GCG Ala	GGA Gly	AAC Asn 95	TGG Trp	TAT Tyr	GGG Gly	TTT Phe	AAC Asn 100	GTT Val	TAC Tyr	CTG Leu	343
Ala	GCG Ala 105	GGC Gly	ATG Met	TTG Leu	CGG Arg	GAA Glu 110	GCC Ala	GCG Ala	GCC Ala	ATG Met	ACC Thr 115	ACA Thr	CAG Gln	ATT Ile	CAG Gln	391
GGC Gly 120	GAT Asp	GTC Val	ATT Ile	CCG Pro	TCC Ser 125	AAT Asn	GTG Val	CCC Pro	GGT Gly	AGC Ser 130	TTT Phe	GCC Ala	ATG Met	GCG Ala	GTT Val 135	439
CGA Arg	CAG Gln	CCA Pro	TGT Cys	GGC Gly 140	Val	GTG Val	CTC Leu	GGT Gly	ATT Ile 145	Ala	CCT Pro	TGG Trp	AAT Asn	GCT Ala 150	Pro	487
GTA Val	ATC Ile	CTT Leu	GGC Gly 155	Val	CGG Arg	GCT Ala	GTT Val	GCG Ala 160	Met	CCG Pro	TTG Leu	GCA Ala	TGC Cys 165	Gly	AAT Asn	535
ACC Thr	GTG Val	GTG Val	Leu	AAA Lys	AGC Ser	TCT Ser	GAG Glu 175	Leu	AGT Ser	CCC Pro	TTT Phe	ACC Thr 180	CAT His	CGC Arg	CTG Leu	583
ATT Ile	GGT Gly 185	Glr	GTG Nal	TTG Leu	CAT His	GAT Asp 190	Ala	GGT	CTG Leu	GGG Gly	GAT Asp 195	Gly	GTG Val	GTC Val	AAT Asn	631
GTC Val 200	ATC	AGC Sei	C AAT	GCC Ala	CCG Pro 205	Glr	A GAC	GCT Ala	CCT Pro	GCG Ala 210	val	GTG Val	GAG Glu	G CGF	CTG Leu 215	679

		·	
			•

ATT Ile	GCA Ala	AAT Asn	CCT Pro	GCG Ala 220	GTA Val	CGT Arg	CGA Arg	GTG Val	AAC Asn 225	TTC Phe	ACC Thr	GGT Gly	TCG Ser	ACC Thr 230	CAC His	727
GTT Val	GGA Gly	CGG Arg	ATC Ile 235	ATT Ile	GGT Gly	GAG Glu	CTG Leu	TCT Ser 240	GCG Ala	CGT Arg	CAT His	CTG Leu	AAG Lys 245	CCT Pro	GCT Ala	775
GTG Val	CTG Leu	GAA Glu 250	TTA Leu	GGT Gly	GGT Gly	AAG Lys	GCT Ala 255	CCG Pro	TTC Phe	TTG Leu	GTC Val	TTG Leu 260	GAC Asp	GAT Asp	GCC Ala	823
GAC Asp	CTC Leu 265	GAT Asp	GCG Ala	GCG Ala	GTC Val	GAA Glu 270	GCG Ala	GCG Ala	GCC Ala	TTT Phe	GGT Gly 275	GCC Ala	TAC Tyr	TTC Phe	AAT Asn	871
CAG Gln 280	GGT Gly	CAA Gln	ATC Ile	TGC Cys	ATG Met 285	TCC Ser	ACT Thr	GAG Glu	CGT Arg	CTG Leu 290	ATT Ile	GTG Val	ACA Thr	GCA Ala	GTC Val 295	919
GCA Ala	GAC Asp	GCC Ala	TTT Phe	GTT Val 300	Glu	AAG Lys	CTG Leu	GCG Ala	AGG Arg 305	AAG Lys	GTC Val	GCC Ala	ACA Thr	CTG Leu 310	CGT Arg	967
GCT Ala	GGC Gly	GAT Asp	CCT Pro 315	Asn	GAT Asp	CCG Pro	CAA Gln	TCG Ser 320	Val	TTG Leu	GGT Gly	TCG Ser	TTG Leu 325	ATT Ile	GAT Asp	1015
GCC Ala	AAT Asn	GCA Ala 330	GGT Gly	CAA Gln	CGC Arg	ATC Ile	CAG Gln 335	Val	CTG Leu	GTC Val	GAT Asp	GAT Asp 340	Ala	CTC Leu	GCA Ala	1063
	Gly	GCG Ala 346	1	TGGP	AA TT Le 42	u Al	C CA	G CG n Ar	C GT g Va	C GA 1 As 42	p Se	G GG r Gl	C AT	T TG e Cy	C CAT s His 430	1113
ATC Ile	C AAT e Asr	GGA Gly	A CCG	ACT Thi 435	: Val	CAT His	GAC Asp	GAG Glu	G GCT 1 Ala 440	Glr	ATC Met	CCA Pro	TTC Phe	GGT Gly 445	GGG Gly	1161
GT( Va]	AAC L Lys	TCC Sei	2 AG0 2 Se1 450	Gly	С ТАС у Туі	GGC Gly	AGC Sei	TTC Phe 455	e Gly	C AGT y Sei	CGA Arg	A GC <i>I</i> g Ala	A TCC a Ser 460	: Ile	GAG Glu	1209
CAC His	C TT'	T ACC e Th: 46	r Glı	G CTO	G CG0 u Arg	TGC Tr	G CTO Let 47	ı Th	C ATT	r CAG e Gli	AA: n Ası	r GGG n Gly 47	y Pro	G CGC	G CAC His	1257
	r Pr	A ATO O Il O 48		TA A	CGAT	CTTC	GGG	CGCC	GCG (	GGCA'	TCAT	GC C	CGCGG	GCGC'	r	1309
CG	CCTC	ТТТА	CAA	тстс	AAT	CTTG	AATA	AA A	CAGA	GCTG	т тс	TCCG	GTCT	TGG	rggatca	1369
AG	GCCA	GTCG	CGG	AGAG	TCT	CGAA	GAGG	AG A	GTAC	AGTG	A AC	GCCG.	AGTC	CAC	ATTGCAA	1429

•
•
-
•
•

CCGCAGGCAT	CATCATGCTC	TGCTCAGCCA	CGCTACCGCA	GTGTGTCGAT	TGGTCATCCT	1489
CCGGTTGAGG	TTACGCAAGA	CGCTGGAGGT	ATTGTCCGGA	TGCGTTCTCT	CGAGGCGCTT	1549
CTTCCCTTCC	CGGGTGGAAT	TC				1571
FIG 20:						

•	
	•
	•

GAAT	TCCG	GCG G	TCGG	CGAA	A GI	'TGAT	'GCGC	TGI	ATCO	STGG	TGAA	AGATO	CAA 1	CCAT	GCTGC	60
GTGA	.CGAG	GC C	CACAC										er Al		T TCC ne Ser	112
														TGC Cys		160
														GTC Val		208
														GAT Asp		256
														CTG Leu 75		304
														CCG Pro		352
														CGG Arg		400
														GTC Val		448
														CGG Arg		496
					_									TGG Trp 155		544
				Pro										GCA Ala		592
	CTG Leu 174		ACAG	CAA	GCGA	ACCG	GA A	TTGC	CAGC	T GG	GGCG	СССТ	CTG	GTAA	GGT	648
TGG	GAAG	ccc	TGCA	AAGT	AA A	CTGG	atgg	С ТТ	тстт	GCCG	CCA	AGGA	тст	GATG	GCGCAG	708
GGG	ATCA	AGA	TCTG	ATCA	AG A	GACA	GGAT	g ag	GATC	GTTT	CGC			GAA Glu	CAA Gln	763

	•
	•

GAT Asp 5	GGA Gly	TTG Leu	CAC His	GCA Ala	GGT Gly 10	TCT Ser	CCG Pro	GCC Ala	GCT Ala	TGG Trp 15	GTG Val	GAG Glu	AGG Arg	CTA Leu	TTC Phe 20	811
GGC Gly	TAT Tyr	GAC Asp	TGG Trp	GCA Ala 25	CAA Gln	CAG Gln	ACA Thr	ATC Ile	GGC Gly 30	TGC Cys	TCT Ser	GAT Asp	GCC Ala	GCC Ala 35	GTG Val	859
TTC Phe	CGG Arg	CTG Leu	TCA Ser 40	GCG Ala	CAG Gln	GGG Gly	CGC Arg	CCG Pro 45	GTT Val	CTT Leu	TTT Phe	GTC Val	AAG Lys 50	ACC Thr	GAC Asp	907
CTG Leu	TCC Ser	GGT Gly 55	GCC Ala	CTG Leu	AAT Asn	GAA Glu	CTG Leu 60	CAG Gln	GAC Asp	GAG Glu	GCA Ala	GCG Ala 65	CGG Arg	CTA Leu	TCG Ser	955
TGG Trp	CTG Leu 70	GCC Ala	ACG Thr	ACG Thr	GGC Gly	GTT Val 75	CCT Pro	TGC Cys	GCA Ala	GCT Ala	GTG Val 80	CTC Leu	GAC Asp	GTT Val	GTC Val	1003
Thr 85	Glu	Ala	Gly	Arg	Asp 90	Trp	Leu	Leu	Leu	Gly 95	Glu	Val	Pro	GGG Gly	Gln 100	1051
Asp	Leu	Leu	Ser	Ser 105	His	Leu	Ala	Pro	Ala 110	Glu	Lys	Val	Ser	ATC Ile 115	Met	1099
Ala	Asp	Ala	Met 120	Arg	Arg	Leu	His	Thr 125	Leu	Asp	Pro	Ala	Thr 130	Cys		1147
Phe	Asp	His 135	Gln	Ala	Lys	His	Arg 140	Ile	Glu	Arg	Ala	145	Thr	Arg		1195
Glu	150	Gly	/ Lev	ı Val	. Asp	Gln 155	Asp	Asp	Leu	Asp	160	ı Glu	n His	: Gln	GGG	1243
Leu 165	Ala	a Pro	Ala	a Glu	170	Ph∈	Ala	Arç	g Leu	175	Ala	a Arg	g Met	: Pro	GAC Asp 180	1291
Gl	/ Glu	Ası د	o Lei	Val 185	l Val	. Thr	His	s Gly	/ Asp 190	Ala	а Су:	s Lei	ı Pro	195		1339
Met	. Va	1 G1	u As: 20	n Gly O	y Ar	g Phe	e Sei	20!	y Ph∈ 5	e Il€	e Ası	p Cy:	s Gl:	y Aro	G CTG	1387
GG' G1	r GT y Va	G GC 1 Al 21	a As	C CGG P Ar	C TA'	r Gl	G GAG n Ası 22	o Il	A GCO	a Le	u Al	T AC a Th 22	r Ar	r GA' g Ası	r ATT o Ile	1435

•	
•	
- <b>v</b>	

GCT GAA GAG CTT GGC GGC GAA TGG GCT GAC CGC TTC CTC GTG CTT TAC Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe Leu Val Leu Tyr 230 235 240	1483
GGT ATC GCC GCT CCC GAT TCG CAG CGC ATC GCC TTC TAT CGC CTT CTT Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe Tyr Arg Leu Leu 255 260	1531
GAC GAG TTC TTC TGA GCGGGACTCT GGGGTTCGAA ATGACCGACC AAGCGACGCC Asp Glu Phe Phe 264	1586
CA TTG AGG GCG CAA GAG GAG AAA TGG ATT GAC CAA GAG ATC GTG GCT Leu Arg Ala Gln Glu Glu Lys Trp Ile Asp Gln Glu Ile Val Ala 197 200 205 210	1633
GTT ACG GAT GAA CAG TTC GAT TTA GAG GGC TAC AAC AGT CGA GCA ATT Val Thr Asp Glu Gln Phe Asp Leu Glu Gly Tyr Asn Ser Arg Ala Ile 215 220 225	1681
GAA CTG CCT CGG AAG GCA AAA TTG TTG ATC GTG ACA GTC ATC CGC GGC Glu Leu Pro Arg Lys Ala Lys Leu Leu Ile Val Thr Val Ile Arg Gly 230 235 240	1729
CTA GCA GTC TTT GAA GCC CTT TCC CGA TTG AAG CCT GTT CAT TCT GGC Leu Ala Val Phe Glu Ala Leu Ser Arg Leu Lys Pro Val His Ser Gly 245 250 255	1777
GGG GTG CAG ACT GCG GGC AAC AGC TGT GCC GTA GTG GAC GGC GCC GCG Gly Val Gln Thr Ala Gly Asn Ser Cys Ala Val Val Asp Gly Ala Ala 260 275	1825
GCG GCT TTG GTG GCT CGA GAG TCG TCT GCG ACA CAG CCG GTC TTG GCT Ala Ala Leu Val Ala Arg Glu Ser Ser Ala Thr Gln Pro Val Leu Ala 280 285 290	1873
AGG ATA CTG GCT ACC TCC GTA GTC GGG ATC GAG CCC GAG CAT ATG GGG Arg Ile Leu Ala Thr Ser Val Val Gly Ile Glu Pro Glu His Met Gly 295 300 305	1921
CTC GGC CCT GCG CCC GCG ATT CGC CTG CTT GCG CGT AGT GAT CTT Leu Gly Pro Ala Pro Ala Ile Arg Leu Leu Leu Ala Arg Ser Asp Leu 310 315 320	1969
AGT TTG AGG GAT ATC GAC CTC TTT GAG ATA AAC GAG GCG CAG GCC GCC Ser Leu Arg Asp Ile Asp Leu Phe Glu Ile Asn Glu Ala Gln Ala Ala 325 330 335	2017
CAA GTT CTA GCG GTA CAG CAT GAA TTG GGT ATT GAG CAC TCA AAA CTT GIn Val Leu Ala Val Gln His Glu Leu Gly Ile Glu His Ser Lys Leu 340 345 350 355	2065
AAT ATT TGG GGC GGG GCC ATT GCA CTT GGA CAC CCG CTT GCC GCG ACC Asn Ile Trp Gly Ala Ile Ala Leu Gly His Pro Leu Ala Ala Thr 360 365 370	2113

		•,
		~ <b>f</b> ;
		•
		•

GGA TTG CGT CTC TGC ATG ACC CTC GCT CAC CAA TTG CAA GCT AAT AAC Gly Leu Arg Leu Cys Met Thr Leu Ala His Gln Leu Gln Ala Asn Asn 375	2161
TTT CGA TAT GGA ATT GCC TCG GCA TGC ATT GGT GGG GGA CAG GGG ATG Phe Arg Tyr Gly Ile Ala Ser Ala Cys Ile Gly Gly Gly Gln Gly Met 390 395 400	2209
GCG GTT CTT TTA GAG AAT CCC CAC TTC GGT TCG TCC TCT GCA CGA AGT Ala Val Leu Leu Glu Asn Pro His Phe Gly Ser Ser Ala Arg Ser 405 410 415	2257
TCG ATG ATT AAC AGA GTT GAC CAC TAT CCA CTG AGC TAA CGGGCATCTC Ser Met Ile Asn Arg Val Asp His Tyr Pro Leu Ser 420 425 430 431	2306
CTTTGTTGCT TTGAGGTGGC GCACGAAGGA GGGCTCGAAA ATCTCTGCTA AAAACAAGAA	2366
GAAGGAACAG GGAACATGAT TAGTTTCGCT CGTATGGCAG AAAGTTTAGG AGTCCAGGCT	2426
AAACTTGCCC TTGCCTTCGC ACTCGTATTA TGTGTCGGGC TGATTGTTAC CGGCACGGGT	2486
TTCTACAGTG TACATACCTT GTCAGGGTTG GTGGGAATTC	2526
FIG. 2p:	

•
*
1

GAAT	TCCG	CG G	TCGG	CGAA	A GT	TGAT	GCGC	TGT	ATCG	TGG '	TGAA	GATC	AA T	CCAT	GCTGC	6	0
GTGA	CGAG	GC C	ACAC	T GT Me	G AG t Se 1	Т <b>Т</b> G r Тr	G TC p Se	A GG r Gl	G GG y G1 5	G GC y Al	Т ТА а Ту	C TC r Se	r Al	G TT a Ph O	T TCC e Ser	11	.2
GAC Asp	ACT Thr	GCG Ala 15	TTG Leu	GTT Val	GCG Ala	GCA Ala	GTG Val 20	CGC Arg	ACC Thr	CCC Pro	TGG Trp	ATT Ile 25	GAT Asp	TGC Cys	GGG Gly	16	50
GGT Gly	GCC Ala 30	CTG Leu	TCG Ser	CTG Leu	GTG Val	TCG Ser 35	CCT Pro	ATC Ile	GAC Asp	TTA Leu	GGG Gly 40	GTA Val	AAG Lys	GTC Val	GCT Ala	20	8
CGC Arg 45	GAA Glu	GTT Val	CTG Leu	ATG Met	CGT Arg 50	GCG Ala	TCG Ser	CTT Leu	GAA Glu	CCA Pro 55	CAA Gln	ATG Met	GTC Val	GAT Asp	AGC Ser 60	25	56
GTA Val	CTC Leu	GCA Ala	GGC Gly	TCT Ser 65	ATG Met	GCT Ala	CAA Gln	GCA Ala	AGC Ser 70	TTT Phe	GAT Asp	GCT Ala	TAC Tyr	CTG Leu 75	CTC Leu	30	04
CCG Pro	CGG Arg	CAC His	ATT Ile 80	GGC Gly	TTG Leu	TAC Tyr	AGC Ser	GGT Gly 85	GTT Val	CCC Pro	AAG Lys	TCG Ser	GTT Val 90	CCG Pro	GCC Ala	3	52
TTG Leu	GGG Gly	GTG Val 95	Gln	CGC Arg	ATT Ile	TGC Cys	GGC Gly 100	ACA Thr	GGC Gly	TTC Phe	GAA Glu	CTG Leu 105	CTT Leu	CGG Arg	CAG Gln	4	00
GCC Ala	GGC Gly 110	Glu	CAG Gln	ATT Ile	TCC Ser	CAA Gln 115	GGC Gly	GCT Ala	GAT Asp	CAC His	GTG Val 120	CTG Leu	TGT Cys	GTC Val	GCG Ala	4	48
GCA Ala 125	Glu	TCC Ser	ATG Met	TCG Ser	CGT Arg 130	AAC Asn	CCC Pro	ATC Ile	GCG Ala	TCG Ser 135	TAT Tyr	ACA Thr	CAC His	CGG Arg	GGC Gly 140	4	96
GGG Gly	TTC Phe	CGC Arg	CTC Leu	GGT Gly 145	Ala	CCC Pro	GTT Val	GAG Glu	TTC Phe 150	Lys	GAT Asp	TTT Phe	TTG Leu	TGG Trp 155	Glu	5	44
GCA Ala	TTC Leu	TTI Phe	GAT Asp	Pro	GCT Ala	CCA Pro	GGA Gly	CTC Lev 165	ı Asp	ATG Met	ATC Ile	GCT Ala	ACC Thr 170	Ala	GAA Glu	5	92
	CTC Let 174	ג	GGAGA	AGGC	GGT	TGCG	TA T	TGGC	GCGCA	AT GC	АТА:	AAAC	TGT	TGTA	ATT	•	648
CAT	AATT	GCAT	TCT	GCCG?	ACA '	rggaa	AGCCE	AT CA	ACAAA	ACGGC	ATO	SATGA	ACC	TGA	ATCGCCA	-	708
GC	GGCA'	rcag	CAC	CTTG'	rcg (	сстто	GCGTA	A TA	TAT	rtgcc	CAT	rggac	CGCA	CACC	CGTGGAA		768
AC	GGAT	GAAG	GCA	CGAA	ccc i	AGTTO	GACAT	ra Ao	GCCT	STTC	GT:	rcgt <i>i</i>	AAAC	TGTA	AATGCAA	, :	828
GT	AGCG	TATG	CGC'	TCAC	GCA .	ACTG	STCC	AG A	ACCT	rgaco	GA/	ACGC <i>I</i>	AGCG	GTG	GTAACGG	}	888

CGCF	GTGG	icg e	STTTI	CATG	G CI	TGTT	TATGA	A CTO	STTTI	TTT	GTAC	CAGTO	TA 1	'GCC'I	'CGGGC	948
ATCC	CAAGO	AGC	CAAGO	GCG	TTAC	CGCCG	STG (	GTC	SATGI	TTG	ATGI	TATO	GA C	CAGO	CAACG	1007
												AGT Ser				1055
												ACA Thr				1103
												GAT Asp 45				1151
												CAG Gln				1199
												ATC Ile				1247
												GCT Ala				1295
												TAT Tyr				1343
												GCG Ala 125				1391
												GTG Val				1439
	Ala										_	TAT Tyr				1487
					Val							CCA Pro				1535
ACC Thr 177		CAA	TTCG	TTC	AAGC	CGAG	AT C	GGCT	TCCC		u Ar			n Gl	G GAG u Glu	1589
			Asp	-				Ala				GAA Glu 215	Gln			1637

	•

TTA Leu	GAG Glu 220	GGC Gly	TAC Tyr	AAC Asn	AGT Ser	CGA Arg 225	GCA Ala	ATT Ile	GAA Glu	CTG Leu	CCT Pro 230	CGG Arg	AAG Lys	GCA Ala	AAA Lys	1685
TTG Leu 235	TTG Leu	ATC Ile	GTG Val	ACA Thr	GTC Val 240	ATC Ile	CGC Arg	GGC Gly	CTA Leu	GCA Ala 245	GTC Val	TTT Phe	GAA Glu	GCC Ala	CTT Leu 250	1733
TCC Ser	CGA Arg	TTG Leu	AAG Lys	CCT Pro 255	GTT Val	CAT His	TCT Ser	GGC Gly	GGG Gly 260	GTG Val	CAG Gln	ACT Thr	GCG Ala	GGC Gly 265	AAC Asn	1,781
AGC Ser	TGT Cys	GCC Ala	GTA Val 270	GTG Val	GAC Asp	GGC Gly	GCC Ala	GCG Ala 275	GCG Ala	GCT Ala	TTG Leu	GTG Val	GCT Ala 280	CGA Arg	GAG Glu	1829
TCG Ser	TCT Ser	GCG Ala 285	ACA Thr	CAG Gln	CCG Pro	GTC Val	TTG Leu 290	GCT Ala	AGG Arg	ATA Ile	CTG Leu	GCT Ala 295	ACC Thr	TCC Ser	GTA Val	1877
GTC Val	GGG Gly 300	ATC Ile	GAG Glu	CCC Pro	GAG Glu	CAT His 305	ATG Met	GGG Gly	CTC Leu	GGC Gly	CCT Pro 310	GCG Ala	CCC Pro	GCG Ala	ATT Ile	1925
Arg 315	Leu	Leu	CTT Leu	Ala	Arg 320	Ser	Asp	Leu	Ser	Leu 325	Arg	Asp	Ile	Asp	Leu 330	1973
Phe	Glu	Ile	AAC Asn	Glu 335	Ala	Gln	Ala	Ala	Gln 340	Val	Leu	Ala	Val	Gln 345	His	2021
GAA Glu	TTG Leu	GGT	ATT Ile 350	Glu	CAC His	TCA Ser	AAA Lys	CTT Leu 355	Asn	ATT	TGG Trp	GGC Gly	GGG Gly 360	Ala	ATT	2069
Ala	Lev	Gly 365	/ His	Pro	Leu	Ala	Ala 370	Thr	Gly	Leu	Arg	1 Leu 375	Cys	Met	ACC Thr	2117
Lev	a Ala 380	n His	s Glr	Leu	Glr	385	Asr	Asr	n Ph∈	e Arg	390	. Gly	, Ile	e Ala	TCG Ser	2165
GC/ Ala 395	a Cys	C ATT	r GGI e Gly	GGC Gly	G GG <i>F</i> 7 Gl <sub>3</sub> 400	/ Glr	G GGC	ATO Met	G GCC	GTT a Val 405	Let	TTA Lei	GAC	AA1 AST	CCC Pro 410	2213
CAC Hi	C TTO	C GG' e Gl	T TCC y Sei	TCC Ser 415	s Se	r GCA	A CGA	A AG	r TCC r Sea 420	r Met	G ATT	AA( e Ası	AGA Arg	425 425	r GAC l Asp 5	2261
			A CTO O Let 43		r	A CG	GGCA'	TCTC	CTT	rgtt(	GCT '	TTGAG	GTG(	GC		2309

			•
			•
	ú.		
			•
			•

GCACGAAGGA	GGGCTCGAAA	ATCTCTGCTA	AAAACAAGAA	GAAGGAACAG	GGAACATGAT	2369
TAGTTTCGCT	CGTATGGCAG	AAAGTTTAGG	AGTCCAGGCT	AAACTTGCCC	TTGCCTTCGC	2429
ACTCGTATTA	TGTGTCGGGC	TGATTGTTAC	CGGCACGGGT	TTCTACAGTG	TACATACCTT	2489
GTCAGGGTTG	GTGGGAATTC					2509
FIG. 2q:						

49

PCT/EP99/07952

WO 00/26355

	_	
		-
		•

GAAT	TCCG	CG G	STCGG	CGAA	A GT	TGAT	GCGC	TGT	ATCG	TGG	TGAA	GATC	аа т	CCAT	GCTGC	60
GTGA	CGAG	GC C	CACAC										r Al		T TCC e Ser	112
			TTG Leu													160
			TCG Ser													208
			CTG Leu													256
			GGC Gly													304
			ATT Ile 80													352
			CAG Gln													400
			CAG Gln													448
			ATG Met													496
			CTC Leu													544
			GAT Asp 160	Pro												592
		Ala	G CGC Arg 5 176	L			la G					rp I			AA GAG ln Glu	641
		Ala	T GTT a Val				Gln					Gly				689
	Ala		r GAA e Glu			Arg					Leu					<b>7</b> 37

	•
	2
	•

													-	
ATC CGC G														785
CAT TCT G														833
GGC GCC G														881
GTC TTG C Val Leu A 290														929
CAT ATG C His Met C 305														977
AGT GAT ( Ser Asp 1														1025
CAG GCC G Gln Ala		Val												1073
TCA AAA G Ser Lys														1121
GCC GCG Ala Ala 370										His				1169
GCT AAT Ala Asn 385									Cys					1217
CAG GGG Gln Gly								His					Ser	1265
GCA CGA Ala Arg		Met					Asp				Leu			1313
CGGGCATC	TC CTT	rgttg	CT T	TGAG	GTGG	ic go	ACGA	AGGA	GGG	CTCG	AAA	ATCT	CTGCTA	1373
AAAACAAG	SAA GAA	GGAAC	AG G	GAAC	ATGA	T TA	GTTI	CGCT	CGI	'ATGG	CAG	AAAG	TTTAGG	1433
AGTCCAGG	SCT AAAG	CTTGC	CC T	TGCC	ттсс	C AC	TCGI	'ATTA	TGI	GTCG	GGC	TGAT	TGTTAC	1493
CGGCACGG	GT TTC	racac	TG T	ACAI	COOR	T GI	'CAGG	GTTG	GTC	GGAA	TTC			1543
FIG. 2r:	:													

•
•
•
•

CTGCAGCCAG	GGCTGAAAAG			GAAGGGAGGG	GACGGCGCCT	60
GGCTCCAATT	GCTCGATGGC		AGTGTCTTGG	GCGCGGTCTT	GGAGAGTTCG	120
GCTAGGGAGA	TAAATTTGCT	GGCCATGGTG	GCGGCCCCTG	ATGGGTTGGA	TGATTTTCTG	180
CATTCTGCAT	CATGAAATTC	ATGAAATCAT	CACTTTTCGG	GGGGTGGGTG	CACGGGATTG	240
AAGGTTGCTA	GGAGAGTGCA	TTGCTCGTAA	GCCCAGGAAG	CACGCGGGTT	TCAGGATGGT	300
GCATGGAAAT	GGCATGAGCT	TTGCTGGATA	TGATTAGAGA	CATTAACTAT	TTTGGCGGAA	360
TGGAAGCACG	ATTCCTCGCC	CGGTAGAGCG	GTAACCGCGA	CATTCAGGAC	CGTAAAAAGG	42,0
AAAGAGCATG	CAACTGACCA	ACAAGAAAAT	CGTCGTCACC	GGAGTGTCCT	CCGGTATCGG	480
TGCCGAAACT	GCCCGCGTTC	TGCGCTCTCA	CGGCGCCACA	GTGATTGGCG	TAGATCGCAA	540
CATGCCGAGC	CTGACTCTGG	ATGCTTTCGT	TCAGGCTGAC	CTGAGCCATC	CTGAAGGCAT	600
CGATAAGGCC	ATCGGGACAG	CAAGCGAACC	GGAATTGCCA	GCTGGGGCGC	CCTCTGGTAA	660
GGTTGGGAAG	CCCTGCAAAG	TAAACTGGAT	GGCTTTCTTG	CCGCCAAGGA	TCTGATGGCG	720
CAGGGGATCA	AGATCTGATC	AAGAGACAGG	ATGAGGATCG	TTTCGCATGA	TTGAACAAGA	780
TGGATTGCAC	GCAGGTTCTC	CGGCCGCTTG	GGTGGAGAGG	CTATTCGGCT	ATGACTGGGC	840
	ATCGGCTGCT	CTGATGCCGC	CGTGTTCCGG	CTGTCAGCGC	AGGGGCGCCC	900
GGTTCTTTTT	GTCAAGACCG	ACCTGTCCGG	TGCCCTGAAT	GAACTGCAGG	ACGAGGCAGC	960
GCGGCTATCG	TGGCTGGCCA	CGACGGGCGT	TCCTTGCGCA	GCTGTGCTCG	ACGTTGTCAC	1020
TGAAGCGGGA	AGGGACTGGC	TGCTATTGGG	CGAAGTGCCG	GGGCAGGATC	TCCTGTCATC	1080
TCACCTTGCT	CCTGCCGAGA	AAGTATCCAT	CATGGCTGAT	GCAATGCGGC	GGCTGCATAC	1140
GCTTGATCCG	GCTACCTGCC	CATTCGACCA	CCAAGCGAAA	CATCGCATCG	AGCGAGCACG	1200
TACTCGGATG	GAAGCCGGTC	TTGTCGATCA	GGATGATCTG	GACGAAGAGC	ATCAGGGGCT	1260
CGCGCCAGCC	GAACTGTTCG	CCAGGCTCAA	GGCGCGCATG	CCCGACGGCG	AGGATCTCGT	1320
CGTGACCCAT	GGCGATGCCT	GCTTGCCGAA	TATCATGGTG	GAAAATGGCC	GCTTTTCTGG	1380
ATTCATCGAC	TGTGGCCGGC	TGGGTGTGGC	GGACCGCTAT	CAGGACATAG	CGTTGGCTAC	1440
CCGTGATATT	GCTGAAGAGC	TTGGCGGCGA	ATGGGCTGAC	CGCTTCCTCG	TGCTTTACGG	1500
TATCGCCGCT	CCCGATTCGC	AGCGCATCGC	CTTCTATCGC	CTTCTTGACG	AGTTCTTCTG	1560
AGCGGGACTC	TGGGGTTCGA	AATGACCGAC	CAAGCGACGC	CCTGGCCGCG	GTGATTGCAT	1620
TCATGTGTGC	TGAGGAGTCA	CGTTGGATCA	ACGGCATAAA	TATTCCAGTG	GACGGAGGTT	1680
TGGCATCGAC	CTACGTGTAA	GTTCGTGGAC	GCCCTTTGCA	CGCGCACTAT	ATCTCTATGC	1740
AGCAGCTGAA	AGCAGCTTTG	GTTTTGATCG	GAGGTAGCGG	GCGGAAAGGT	GCAGAATGTC	1800
AATAATAA	AGGATTCTTG	TGAAGCTTTA	GTTGTCCGTA	AACGAAAATA	AAAATAAAGA	1860
GGAATGATAT	GAAAGCAAGT	AGATCAGTCT	GCACTTTCAA	AATAGCTACC	CTGGCAGGCG	1920
CCATTTATGC	AGCGCTGCCA	ATGTCAGCTG	CAAACTCGAT	GCAGCTGGAT	GTAGGTAGCT	1980
CGGATTGGAC	GGTGCGTTGG	GGACAACACC	CTCAAGTATA	GCCTTGCCTC	TCGCCTGAAT	2040
GAGCAAGACT				GTTATATCCG	GATATTCAAA	2100
GTCAGGGTGA	TCGTAACTTT	GACCGGGGGC	TTGGTATCCA	ATCGTCTCGA	TATTCTGGCT	2160
GCAG						2164

	•

CTGCAGCCAG	GGCTGAAAAG				GACGGCGCCT	60
GGCTCCAATT	GCTCGATGGC	GCCGCGATTG	AGTGTCTTGG	GCGCGGTCTT	GGAGAGTTCG	120
GCTAGGGAGA	TAAATTTGCT	GGCCATGGTG	GCGGCCCCTG	ATGGGTTGGA	TGATTTTCTG	180
CATTCTGCAT	CATGAAATTC	ATGAAATCAT	CACTTTTCGG	GGGGTGGGTG	CACGGGATTG	240
AAGGTTGCTA	GGAGAGTGCA	TTGCTCGTAA	GCCCAGGAAG	CACGCGGGTT	TCAGGATGGT	300
GCATGGAAAT	GGCATGAGCT	TTGCTGGATA	TGATTAGAGA	CATTAACTAT	TTTGGCGGAA	360
TGGAAGCACG	ATTCCTCGCC	CGGTAGAGCG	GTAACCGCGA	CATTCAGGAC	CGTAAAAAGG	420
AAAGAGCATG	CAACTGACCA	ACAAGAAAAT	CGTCGTCACC	GGAGTGTCCT	CCGGTATCGG	480
TGCCGAAACT	GCCCGCGTTC	TGCGCTCTCA	CGGCGCCACA	GTGATTGGCG	TAGATCGCAA	540
CATGCCGAGC	CTGACTCTGG	ATGCTTTCGT	TCAGGCTGAC	CTGAGCCATC	CTGAGGGGAG	600
AGGCGGTTTG	CGTATTGGGC	GCATGCATAA	AAACTGTTGT	AATTCATTAA	GCATTCTGCC	660
GACATGGAAG	CCATCACAAA	CGGCATGATG	AACCTGAATC	GCCAGCGGCA	TCAGCACCTT	720
GTCGCCTTGC	GTATAATATT	TGCCCATGGA	CGCACACCGT	GGAAACGGAT	GAAGGCACGA	780
ACCCAGTTGA	CATAAGCCTG	TTCGGTTCGT	AAACTGTAAT	GCAAGTAGCG	TATGCGCTCA	840
CGCAACTGGT	CCAGAACCTT	GACCGAACGC	AGCGGTGGTA	ACGGCGCAGT	GGCGGTTTTC	900
ATGGCTTGTT	ATGACTGTTT	TTTTGTACAG	TCTATGCCTC	GGGCATCCAA	GCAGCAAGCG	960
CGTTACGCCG	TGGGTCGATG	TTTGATGTTA	TGGAGCAGCA	ACGATGTTAC	GCAGCAGCAA	1020
CGATGTTACG	CAGCAGGGCA	GTCGCCCTAA	AACAAAGTTA	GGTGGCTCAA	GTATGGGCAT	1080
CATTCGCACA	TGTAGGCTCG	GCCCTGACCA	AGTCAAATCC	ATGCGGGCTG	CTCTTGATCT	1140
TTTCGGTCGT	GAGTTCGGAG	ACGTAGCCAC	CTACTCCCAA	CATCAGCCGG	ACTCCGATTA	1200
CCTCGGGAAC	TTGCTCCGTA	GTAAGACATT	CATCGCGCTT	GCTGCCTTCG	ACCAAGAAGC	1260
GGTTGTTGGC	GCTCTCGCGG	CTTACGTTCT	GCCCAGGTTT	GAGCAGCCGC	GTAGTGAGAT	1320
CTATATCTAT	GATCTCGCAG	TCTCCGGCGA	GCACCGGAGG	CAGGGCATTG	CCACCGCGCT	1380
CATCAATCTC	CTCAAGCATG	AGGCCAACGC	GCTTGGTGCT	TATGTGATCT	ACGTGCAAGC	1440
AGATTACGGT	GACGATCCCG	CAGTGGČTCT	CTATACAAAG	TTGGGCATAC	GGGAAGAAGT	1500
GATGCACTTT	GATATCGACC	CAAGTACCGC	CACCTAACAA	TTCGTTCAAG	CCGAGATCGG	1560
CTTCCCTGAT	TGCATTCATG	TGTGCTGAGG	AGTCACGTTG	GATCAACGGC	ATAAATATTC	1620
CAGTGGACGG	AGGTTTGGCA	TCGACCTACG	TGTAAGTTCG	TGGACGCCCT	TTGCACGCGC	1680
ACTATATCTC	TATGCAGCAG	CTGAAAGCAG	CTTTGGTTTT	GATCGGAGGT	AGCGGGCGGA	1740
AAGGTGCAGA	ATGTCTAAAT	AATAAAGGAT	TCTTGTGAAG	CTTTAGTTGT	CCGTAAACGA	1800
AAATAAAAT	AAAGAGGAAT	GATATGAAAG	CAAGTAGATC	AGTCTGCACT	TTCAAAATAG	1860
CTACCCTGGC	AGGCGCCATT	TATGCAGCGC	TGCCAATGTC	AGCTGCAAAC	TCGATGCAGC	1920
TGGATGTAGG	TAGCTCGGAT	TGGACGGTGC	GTTGGGGACA	ACACCCTCAA	GTATAGCCTT	1980
GCCTCTCGCC	TGAATGAGCA	AGACTCAAGT	CTGACAAATG	CGCCGACTGT	CAATGGTTAT	2040
ATCCGGATAT	TCAAAGTCAG	GGTGATCGTA	ACTTTGACCG	GGGGCTTGGT	ATCCAATCGT	2100
CTCGATATTC	TGGCTGCAG					2119

· ·
•
$\delta$

CTGCAGCCAG	GGCTGAAAAG	GAGGGATTCA	GTGAGGTCAT	GAAGGGAGGG	GACGGCGCCT	60
GGCTCCAATT	GCTCGATGGC	GCCGCGATTG	AGTGTCTTGG	GCGCGGTCTT	GGAGAGTTCG	120
GCTAGGGAGA	TAAATTTGCT	GGCCATGGTG	GCGGCCCCTG	ATGGGTTGGA	TGATTTTCTG	180
CATTCTGCAT	CATGAAATTC	ATGAAATCAT	CACTTTTCGG	GGGGTGGGTG	CACGGGATTG	240
AAGGTTGCTA	GGAGAGTGCA	TTGCTCGTAA	GCCCAGGAAG	CACGCGGGTT	TCAGGATGGT	300
GCATGGAAAT	GGCATGAGCT	TTGCTGGATA	TGATTAGAGA	CATTAACTAT	TTTGGCGGAA	360
TGGAAGCACG	ATTCCTCGCC	CGGTAGAGCG	GTAACCGCGA	CATTCAGGAC	CGTAAAAAGG	420
AAAGAGCATG	CAACTGACCA	ACAAGAAAAT	CGTCGTCACC	GGAGTGTCCT	CCGGTATCGG	480
TGCCGAAACT	GCCCGCGTTC	TGCGCTCTCA	CGGCGCCACA	GTGATTGGCG	TAGATCGCAA	540
CATGCCGAGC	CTGACTCTGG	ATGCTTTCGT	TCAGGCTGAC	CTGAGCCATC	CTGAAGGCAT	600
CGATCAACGG	CATAAATATT	CCAGTGGACG	GAGGTTTGGC	ATCGACCTAC	GTGTAAGTTC	660
GTGGACGCCC	TTTGCACGCG	CACTATATCT	CTATGCAGCA	GCTGAAAGCA	GCTTTGGTTT	720
TGATCGGAGG	TAGCGGGCGG	AAAGGTGCAG	AATGTCTAAA	TAATAAAGGA	TTCTTGTGAA	780
GCTTTAGTTG	TCCGTAAACG	AAAATAAAA	TAAAGAGGAA	TGATATGAAA	GCAAGTAGAT	840
CAGTCTGCAC	TTTCAAAATA	GCTACCCTGG	CAGGCGCCAT	TTATGCAGCG	CTGCCAATGT	900
CAGCTGCAAA	CTCGATGCAG	CTGGATGTAG	GTAGCTCGGA	TTGGACGGTG	CGTTGGGGAC	960
AACACCCTCA	AGTATAGCCT	TGCCTCTCGC	CTGAATGAGC	AAGACTCAAG	TCTGACAAAT	1020
GCGCCGACTG	TCAATGGTTA	TATCCGGATA	TTCAAAGTCA	GGGTGATCGT	AACTTTGACC	1080
GGGGGCTTGG	TATCCAATCG	TCTCGATATT	CTGGCTGCAG			1120

		_
		•
		•
		*

GAATTCCGCG	TATCGCCCGG	TTCTATCAGC	GGGCCGCTTT	CGAAAGTCAT	GGTGTTAGCC	60
GGTAGGGTCT	TTTTCTTGGC	CATGCTTGTT	GCCTGAACCT	TCGTTGACAT	AGGGCAGAGG	120
TGCGTTTGCC	GCTTCGCTTC	GCGATGAACC	GCATCGAGAT	GCTGAGGTCA	GGATTTTTCC	180
TTAACTCGCG	TAAGCATTCT	GTCATTTTTT	TGGTGGCTTT	GAACAGCCTG	ATGAAAGGTG	240
GTCTCGCCCT	TTGAGGCCGA	TTCTTGGGCG	CTTGGCGGCG	TCGAAGCGAT	GCTCCACTAC	300
CGATTAAGAT	ATAAAATAA	AGGAAACCGC	ATGGTTTCTT	ATGTGAATTT	GTCTGGCATA	360
CTCCAGCTCA	AGGGCAATTT	TTGGGCTATT	GGCTGAGCAG	TTGCCTCTAT	ATGGTTATTC	420
AGAATAACAA	TTGACTCCTC	AGGAGGTCAG	CGATGAGCAT	TCTTGGTTTG	AATGGTGCCC.	480
CGGTCGGAGC	TGAGCAGCTG	GGCTCGGCTC	TTGATCGCAT	GAAGAAGGCG	CACCTGGAGC	540
AGGGGCCTGC	AAACTTGGAG	CTGCGTCTGA	GTAGGCTGGA	TCGTGCGATT	GCAATGCTTC	600
TGGAAAATCG	TGAAGCAATT	GCCGACGCGG	TTTCTGCTGA	CTTTGGCAAT	CGCAGCCGTG	660
AGCAAACACT	GCTTTGCGAC	ATTGCTGGCT	CGGTGGCAAG	CCTGAAGGAT	AGCCGCGAGC	720
ACGTGGCCAA	ATGGATGGAG	CCCGAACATC	ACAAGGCGAT	GTTTCCAGGG	GCGGAGGCAC	780
GCGTTGAGTT	TCAGCCGCTG	GGTGTCGTTG	GGGTCATTAG	TCCCTGGAAC	TTCCCTATCG	840
TACTGGCCTT	TGGGCCGCTG	GCCGGCATAT	TCGCAGCAGG	TAATCGCGCC	ATGCTCAAGC	900
CGTCCGAGCT	TACCCCGCGG	ACTTCTGCCC	TGCTTGCGGA	GCTAATTGCT	CGTTACTTCG	960
ATGAAACTGA	GCTGACTACA	GTGCTGGGCG	ACGCTGAAGT	CGGTGCGCTG	TTCAGTGCTC	1020
AGCCTTTCGA	TCATCTGATC	TTCACCGGCG	GCACTGCCGT	GGCCAAGCAC	ATCATGCGTG	1080
CCGCGGCGGA	TAACCTAGTG	CCCGTTACCC	TGGAATTGGG	TGGCAAATCG	CCGGTGATCG	1140
TTTCCCGCAG	TGCAGATATG	GCGGACGTTG	CACAACGGGT	GTTGACGGTG	AAAACCTTCA	1200
ATGCCGGGCA	AATCTGTCTG	GCACCGGACT	ATGTGCTGCT		ACAGCAAGCG	1260
AACCGGAATT	GCCAGCTGGG	GCGCCCTCTG	GTAAGGTTGG	GAAGCCCTGC	AAAGTAAACT	1320
GGATGGCTTT		AGGATCTGAT	GGCGCAGGGG	ATCAAGATCT	GATCAAGAGA	1380
CAGGATGAGG	ATCGTTTCGC	ATGATTGAAC	AAGATGGATT	GCACGCAGGT	TCTCCGGCCG	1440
CTTGGGTGGA	GAGGCTATTC	GGCTATGACT	GGGCACAACA	GACAATCGGC	TGCTCTGATG	1500
CCGCCGTGTT	CCGGCTGTCA	GCGCAGGGGC	GCCCGGTTCT	TTTTGTCAAG	ACCGACCTGT	1560
CCGGTGCCCT	GAATGAACTG	CAGGACGAGG			GCCACGACGG	1620
GCGTTCCTTG	CGCAGCTGTG	CTCGACGTTG			TGGCTGCTAT	1680
TGGGCGAAGT	GCCGGGGCAG	GATCTCCTGT	CATCTCACCT	TGCTCCTGCC	GAGAAAGTAT	1740
CCATCATGGC	TGATGCAATG		ATACGCTTGA		TGCCCATTCG	1800
ACCACCAAGC	GAAACATCGC		CACGTACTCG		GGTCTTGTCG	1860
ATCAGGATGA	TCTGGACGAA	GAGCATCAGG		AGCCGAACTG	TTCGCCAGGC	1920
TCAAGGCGCG	CATGCCCGAC	GGCGAGGATC	TCGTCGTGAC	CCATGGCGAT	GCCTGCTTGC	1980
CGAATATCAT	GGTGGAAAAT	GGCCGCTTTT		CGACTGTGGC		2040
TGGCGGACCG	• • • • • • • • • • • • • • • • • • • •				GAGCTTGGCG	2100
GCGAATGGGC	TGACCGCTTC	CTCGTGCTTT	ACGGTATCGC	CGCTCCCGAT	TCGCAGCGCA	2160

	•
	•
	•
	2
	•

TCGCCTTCTA	TCGCCTTCTT	GACGAGTTCT	TCTGAGCGGG	ACTCTGGGGT	TCGAAATGAC	2220
CGACCAAGCG	ACGCCCGCCA	TGCCAAGCCT	GTTCTCGTGC	AAAGTCCTGT	GGGTGAGTCG	2280
		CTACGGAGAA				2340
		TTGGTAGTGG				2400
		CATTGGATAG				2460
TGATGAGCAC	CCTGGAAGGC	GCGCTGTACG	CGGACGACTG	GGTTCATCTT	CGCCATTCAT	2520
		TACCGCGATG				2580
		GGGGCGGGG				2640
		GCGGCGGCTG				270.0
		GCTGCCGGCG				2760
CGAGCAAGGA	AGAACTCATT	TCCCGGTTAG	TGGCTCGAGA	CATGTCCAAC	CTTGAGGAAT	2820
TC						2822

		•
•		
<i>i</i> /		
		2
		-
		•

GAATTCCGCG		TTCTATCAGC		CGAAAGTCAT	GGTGTTAGCC	60
GGTAGGGTCT	TTTTCTTGGC	CATGCTTGTT	GCCTGAACCT	TCGTTGACAT	AGGGCAGAGG	120
TGCGTTTGCC	GCTTCGCTTC	GCGATGAACC	GCATCGAGAT	GCTGAGGTCA	GGATTTTTCC	180
TTAACTCGCG	TAAGCATTCT	GTCATTTTT	TGGTGGCTTT	GAACAGCCTG	ATGAAAGGTG	240
GTCTCGCCCT	TTGAGGCCGA	TTCTTGGGCG	CTTGGCGGCG	TCGAAGCGAT	GCTCCACTAC	300
CGATTAAGAT	ATTAAAATT	AGGAAACCGC	ATGGTTTCTT	ATGTGAATTT	GTCTGGCATA	360
CTCCAGCTCA	AGGGCAATTT	TTGGGCTATT	GGCTGAGCAG	TTGCCTCTAT	ATGGTTATTC	420
AGAATAACAA		AGGAGGTCAG		TCTTGGTTTG	AATGGTGCCC	480
CGGTCGGAGC		GGCTCGGCTC		GAAGAAGGCG		540
AGGGGCCTGC		CTGCGTCTGA			GCAATGCTTC	600
TGGAAAATCG		GCCGACGCGG			CGCAGCCGTG	660
AGCAAACACT	GCTTTGCGAC	ATTGCTGGCT	CGGTGGCAAG	CCTGAAGGAT	AGCCGCGAGC	720
ACGTGGCCAA	ATGGATGGAG	CCCGAACATC	ACAAGGCGAT	GTTTCCAGGG	GCGGAGGCAC	780
GCGTTGAGTT	TCAGCCGCTG	GGTGTCGTTG	GGGTCATTAG	TCCCTGGAAC	TTCCCTATCG	840
TACTGGCCTT	TGGGCCGCTG	GCCGGCATAT	TCGCAGCAGG	TAATCGCGCC	ATGCTCAAGC	900
CGTCCGAGCT		ACTTCTGCCC	TGCTTGCGGA	GCTAATTGCT	CGTTACTTCG	960
ATGAAACTGA	GCTGACTACA	GTGCTGGGCG	ACGCTGAAGT	CGGTGCGCTG	TTCAGTGCTC	1020
AGCCTTTCGA	TCATCTGATC	TTCACCGGCG	GCACTGCCGT	GGCCAAGCAC	ATCATGCGTG	1080
CCGCGGCGGA			TGGAATTGGG	TGGCAAATCG	CCGGTGATCG	1140
TTTCCCGCAG	TGCAGATATG	GCGGACGTTG	CACAACGGGT	GTTGACGGTG	AAAACCTTCA	1200
ATGCCGGGCA		·	ATGTGCTGGG	GGAGAGGCGG	TTTGCGTATT	1260
GGGCGCATGC	ATAAAAACTG	TTGTAATTCA	TTAAGCATTC	TGCCGACATG	GAAGCCATCA	1320
CAAACGGCAT		AATCGCCAGC			TTGCGTATAA	1380
TATTTGCCCA	TGGACGCACA	CCGTGGAAAC	GGATGAAGGC	ACGAACCCAG	TTGACATAAG	1440
CCTGTTCGGT	TCGTAAACTG	TAATGCAAGT		CTCACGCAAC	TGGTCCAGAA	1500
		GGTAACGGCG		TTTCATGGCT	TGTTATGACT	1560
GTTTTTTTGT		CCTCGGGCAT			GCCGTGGGTC	1620
GATGTTTGAT		AGCAACGATG		GCAACGATGT	TACGCAGCAG	1680
GGCAGTCGCC		GTTAGGTGGC		GCATCATTCG	CACATGTAGG	1740
		ATCCATGCGG			TCGTGAGTTC	1800
		CCAACATCAG		ATTACCTCGG	GAACTTGCTC	1860
		GCTTGCTGCC		AAGCGGTTGT	TGGCGCTCTC	1920
GCGGCTTACG		GTTTGAGCAG		AGATCTATAT	CTATGATCTC	1980
_		GAGGCAGGGC		CGCTCATCAA	TCTCCTCAAG	2040
CATGAGGCCA		TGCTTATGTG		AAGCAGATTA	CGGTGACGAT	2100
CCCGCAGTGG	CTCTCTATAC	AAAGTTGGGC	ATACGGGAAG	AAGTGATGCA	CTTTGATATC	2160

			•
			•
			,
			,
			,
			,
			,
			,
			,
			,
			,
			,

GACCCAAGTA	CCGCCACCTA	ACAATTCGTT	CAAGCCGAGA	TCGGCTTCCC	TGCAAAGTCC	2220
TGTGGGTGAG	TCGAACTTGG	CGATGCGCGC	ACCCTACGGA	GAAGCGATCC	ACGGACTGCT	2280
CTCTGTCCTC	CTTTCAACGG	AGTGTTAGAA	CCGTTGGTAG	TGGTTTTGGA	CGGGCCCAGG	2340
AGCATGCGCT	TCTGGGCCCG	TTTCTTGAGT	ATTCATTGGA	TAGTCACGCG	TGGTAGCTTC	2400
GAGCCTGCAC	AGCTGATGAG	CACCCTGGAA	GGCGCGCTGT	ACGCGGACGA	CTGGGTTCAT	2460
CTTCGCCATT	CATGACGGAA	CTCCGTTCCC	CAGTACCGCG	ATGACTATTT	TGCCTCTTCC	2520
GATGTCCGAT	TCCACGCCGC	CTGACGCTAA	GCGGGGGCGG	GGGCGCCCGC	ATCCCAGCCC	2580
AGACAGCAAC	AAATGAGTAG	GCTCTTGGAT	GCCGCGGCGG	CTGAGATTGG	TAACGGCAAT	2640
TTCGTCAATG	TGACGATGGA	TTCGATTGCC	CGTGCTGCCG	GCGTCTCAAA	AAAAACGCTG	2700
TACGTCTTGG	TGGCGAGCAA	GGAAGAACTC	ATTTCCCGGT	TAGTGGCTCG	AGACATGTCC	2760
AACCTTGAGG	AATTC					2775

•
1
•

GAATTCCGCG	TATCGCCCGG	TTCTATCAGC	GGGCCGCTTT	CGAAAGTCAT	GGTGTTAGCC	60
GGTAGGGTCT	TTTTCTTGGC	CATGCTTGTT	GCCTGAACCT	TCGTTGACAT	AGGGCAGAGG	120
TGCGTTTGCC	GCTTCGCTTC	GCGATGAACC	GCATCGAGAT	GCTGAGGTCA	GGATTTTTCC	180
TTAACTCGCG	TAAGCATTCT	GTCATTTTT	TGGTGGCTTT	GAACAGCCTG	ATGAAAGGTG	240
GTCTCGCCCT	TTGAGGCCGA	TTCTTGGGCG	CTTGGCGGCG	TCGAAGCGAT	GCTCCACTAC	300
CGATTAAGAT	ATTAAAATA	AGGAAACCGC	ATGGTTTCTT	ATGTGAATTT	GTCTGGCATA	360
CTCCAGCTCA	AGGGCAATTT	TTGGGCTATT	GGCTGAGCAG	TTGCCTCTAT	ATGGTTATTC	42.0
AGAATAACAA	TTGACTCCTC	AGGAGGTCAG	CGATGAGCAT	TCTTGGTTTG	AATGGTGCCC	480
CGGTCGGAGC	TGAGCAGCTG	GGCTCGGCTC	TTGATCGCAT	GAAGAAGGCG	CACCTGGAGC	540
AGGGGCCTGC	AAACTTGGAG	CTGCGTCTGA	GTAGGCTGGA	TCGTGCGATT	GCAATGCTTC	600
TGGAAAATCG	TGAAGCAATT	GCCGACGCGG	TTTCTGCTGA	CTTTGGCAAT	CGCAGCCGTG	660
AGCAAACACT	GCTTTGCGAC	ATTGCTGGCT	CGGTGGCAAG	CCTGAAGGAT	AGCCGCGAGC	720
ACGTGGCCAA	ATGGATGGAG	CCCGAACATC	ACAAGGCGAT	GTTTCCAGGG	GCGGAGGCAC	780
GCGTTGAGTT	TCAGCCGCTG	GGTGTCGTTG	GGGTCATTAG	TCCCTGGAAC	TTCCCTATCG	840
TACTGGCCTT	TGGGCCGCTG	GCCGGCATAT	TCGCAGCAGG	TAATCGCGCC	ATGCTCAAGC	900
CGTCCGAGCT	TACCCCGCGG	ACTTCTGCCC	TGCTTGCGGA	GCTAATTGCT	CGTTACTTCG	960
ATGAAACTGA	GCTGACTACA	GTGCTGGGCG	ACGCTGAAGT	CGGTGCGCTG	TTCAGTGCTC	1020
AGCCTTTCGA	TCATCTGATC	TTCACCGGCG	GCACTGCCGT	GGCCAAGCAC	ATCATGCGTG	1080
CCGCGGCGGA	TAACCTAGTG	CCCGTTACCC	TGGAATTGGG	TGGCAAATCG	CCGGTGATCG	1140
TTTCCCGCAG	TGCAGATATG	GCGGACGTTG	CACAACGGGT	GTTGACGGTG	AAAACCTTCA	1200
ATGCCGGGCA	AATCTGTCTG	GCACCGTGGG	TGAGTCGAAC	TTGGCGATGC	GCGCACCCTA	1260
CGGAGAAGCG	ATCCACGGAC	TGCTCTCTGT	CCTCCTTTCA	ACGGAGTGTT	AGAACCGTTG	1320
GTAGTGGTTT	TGGACGGGCC	CAGGAGCATG	CGCTTCTGGG	CCCGTTTCTT	GAGTATTCAT	1380
TGGATAGTCA	CGCGTGGTAG	CTTCGAGCCT	GCACAGCTGA	TGAGCACCCT	GGAAGGCGCG	1440
CTGTACGCGG	ACGACTGGGT	TCATCTTCGC	CATTCATGAC	GGAACTCCGT	TCCCCAGTAC	1500
CGCGATGACT	ATTTTGCCTC	TTCCGATGTC	CGATTCCACG	CCGCCTGACG	CTAAGCGGGG	1560
GCGGGGGCGC	CCGCATCCCA	GCCCAGACAG	CAACAAATGA	GTAGGCTCTT	GGATGCCGCG	1620
GCGGCTGAGA	TTGGTAACGG		AATGTGACGA		TGCCCGTGCT	1680
GCCGGCGTCT	CAAAAAAAAC	GCTGTACGTC		GCAAGGAAGA	ACTCATTTCC	1740
CGGTTAGTGG	CTCGAGACAT	GTCCAACCTT	GAGGAATTC			1779



CTGCAGCCGA		AGCACTTTAC			TTCAGAATGG	60
CCCGCGGCAC		AAATCGATCT		CGGGCATCAT	GCCCGCGGCG	120
CTCGCCTCAT		AACTTGATAA		GTTCTCCGGT	CTTGGTGGAT	180
CAAGGCCAGT	CGCGGAGAGT	CTCGAAGAGG		GAACGCCGAG	TCCACATTGC	240
AACCGCAGGC	ATCATCATGC	TCTGCTCAGC	CACGCTACCG	CAGTGTGTCG	ATTGGTCATC	300
CTCCGGTTGA		GACGCTGGAG		GATGCGTTCT	CTCGAGGCGC	360
TTCTTCCCTT	CCCGGGTCGA	ATTCTTGAGC	GTCTCGAGCA	TTGGGCTAAG	ACCCGTCCAG	420
AACAAACCTG	CGTTGCTGCC	AGGGCGGCAA	ATGGGGAATG	GCGTCGTATC	AGCTACGCGG	480
AAATGTTCCA	CAACGTCCGC	GCCATCGCAC	AGAGCTTGCT	TCCTTACGGA	CTATCGGCAG	540
AGCGTCCGCT	GCTTATCGTC	TCTGGAAATG	ACCTGGAACA	TCTTCAGCTG	GCATTTGGGG	600
CTATGTATGC	GGGCATTCCC	TATTGCCCGG	TGTCTCCTGC	TTATTCACTG	CTGTCGCAAG	660
ATTTGGCGAA	GCTGCGTCAC	ATCGTAGGTC	TTCTGCAACC	GGGACTGGTC	TTTGCTGCCG	720
ATGCAGCACC	TTTCCAGGGG	ACAGCAAGCG	AACCGGAATT	GCCAGCTGGG	GCGCCCTCTG	780
GTAAGGTTGG	GAAGCCCTGC	AAAGTAAACT	GGATGGCTTT	CTTGCCGCCA	AGGATCTGAT	840
GGCGCAGGGG	ATCAAGATCT	GATCAAGAGA	CAGGATGAGG	ATCGTTTCGC	ATGATTGAAC	900
AAGATGGATT	GCACGCAGGT	TCTCCGGCCG	CTTGGGTGGA	GAGGCTATTC	GGCTATGACT	960
GGGCACAACA	GACAATCGGC	TGCTCTGATG	CCGCCGTGTT	CCGGCTGTCA	GCGCAGGGGC	1020
GCCCGGTTCT	TTTTGTCAAG	ACCGACCTGT	CCGGTGCCCT	GAATGAACTG	CAGGACGAGG	1080
CAGCGCGGCT	ATCGTGGCTG	GCCACGACGG	GCGTTCCTTG	CGCAGCTGTG	CTCGACGTTG	1140
TCACTGAAGC	GGGAAGGGAC	TGGCTGCTAT	TGGGCGAAGT	GCCGGGGCAG	GATCTCCTGT	1200
CATCTCACCT	TGCTCCTGCC	GAGAAAGTAT	CCATCATGGC	TGATGCAATG	CGGCGGCTGC	1260
ATACGCTTGA	TCCGGCTACC	TGCCCATTCG	ACCACCAAGC	GAAACATCGC	ATCGAGCGAG	1320
CACGTACTCG	GATGGAAGCC	GGTCTTGTCG	ATCAGGATGA	TCTGGACGAA	GAGCATCAGG	1380
GGCTCGCGCC	AGCCGAACTG	TTCGCCAGGC	TCAAGGCGCG	CATGCCCGAC	GGCGAGGATC	1440
TCGTCGTGAC	CCATGGCGAT	GCCTGCTTGC	CGAATATCAT	GGTGGAAAAT	GGCCGCTTTT	1500
CTGGATTCAT	CGACTGTGGC	CGGCTGGGTG	TGGCGGACCG	CTATCAGGAC	ATAGCGTTGG	1560
CTACCCGTGA	TATTGCTGAA	GAGCTTGGCG	GCGAATGGGC	TGACCGCTTC	CTCGTGCTTT	1620
ACGGTATCGC	CGCTCCCGAT	TCGCAGCGCA	TCGCCTTCTA	TCGCCTTCTT	GACGAGTTCT	1680
TCTGAGCGGG	ACTCTGGGGT	TCGAAATGAC	CGACCAAGCG	ACGCCCCTGT	TTTGCAATGG	1740
CGGTCGGCGA	AAGTTGATGC	GCTGTATCGT	GGTGAAGATC	AATCCATGCT	GCGTGACGAG	1800
GCCACACTGT			ACTCGGCGTT	TTCCGACACT	GCGTTGGTTG	1860
CGGCAGTGCG	CACCCCCTGG	ATTGATTGCG	GGGGTGCCCT	GTCGCTGGTG	TCGCCTATCG	1920
ACTTAGGGGT	AAAGGTCGCT	CGCGAAGTTC	TGATGCGTGC	GTCGCTTGAA	CCACAAATGG	1980
TCGATAGCGT	ACTCGCAGGC	TCTATGGCTC	AAGCAAGCTT	TGATGCTTAC	CTGCTCCCGC	2040
	CTTGTACAGC		AGTCGGTTCC			2100
			AGGCCGGCGA	GCAGATTTCC	CAAGGCGCTG	2160
ATCACGTGCT	GTGTGTCGCG	GGCTGCAG				2188

	•
	•
·	
	,
	,
	,
	,
	•
	,
	,
	,
	•
	•
	,

CMCCACCCCA	GCATCGATTG	AGCACTTTAC	CCAGCTGCGC	TGGCTGACCA	TTCAGAATGG	60
CCCGCGGCAC	TATCCAATCT	AAATCGATCT	TCGGGCGCCG	CGGGCATCAT	GCCCGCGGCG	120
CTCGCCTCAT		AACTTGATAA		GTTCTCCGGT	CTTGGTGGAT	180
CAAGGCCAGT	CGCGGAGAGT	CTCGAAGAGG	AGAGTACAGT	GAACGCCGAG	TCCACATTGC	240
AACCGCAGGC	ATCATCATGC	TCTGCTCAGC	CACGCTACCG	CAGTGTGTCG	ATTGGTCATC	300
CTCCGGTTGA		GACGCTGGAG	GTATTGTCCG	GATGCGTTCT	CTCGAGGCGC	360
TTCTTCCCTT	CCCGGGTCGA	ATTCTTGAGC	GTCTCGAGCA	TTGGGCTAAG	ACCCGTCCAG	420
AACAAACCTG	CGTTGCTGCC	AGGGCGGCAA	ATGGGGAATG	GCGTCGTATC	AGCTACGCGG	480
AAATGTTCCA		GCCATCGCAC	AGAGCTTGCT	TCCTTACGGA	CTATCGGCAG	540
AGCGTCCGCT	GCTTATCGTC	TCTGGAAATG	ACCTGGAACA	TCTTCAGCTG	GCATTTGGGG	600
CTATGTATGC		TATTGCCCGG	TGTCTCCTGC	TTATTCACTG	CTGTCGCAAG	660
	GCTGCGTCAC	ATCGTAGGTC	TTCTGCAACC	GGGACTGGTC	TTTGCTGCCG	720
ATGCAGCACC	TTTCCAGGGG	GAGAGGCGGT	TTGCGTATTG	GGCGCATGCA	TAAAAACTGT	780
TGTAATTCAT	TAAGCATTCT	GCCGACATGG	AAGCCATCAC	AAACGGCATG	ATGAACCTGA	840
ATCGCCAGCG	GCATCAGCAC	CTTGTCGCCT	TGCGTATAAT	ATTTGCCCAT	GGACGCACAC	900
CGTGGAAACG	GATGAAGGCA	CGAACCCAGT	TGACATAAGC	CTGTTCGGTT	CGTAAACTGT	960
AATGCAAGTA	GCGTATGCGC	TCACGCAACT	GGTCCAGAAC	CTTGACCGAA	CGCAGCGGTG	1020
GTAACGGCGC	AGTGGCGGTT	TTCATGGCTT	GTTATGACTG	TTTTTTTGTA	CAGTCTATGC	1080
CTCGGGCATC	CAAGCAGCAA	GCGCGTTACG	CCGTGGGTCG	ATGTTTGATG	TTATGGAGCA	1140
GCAACGATGT	TACGCAGCAG	CAACGATGTT	ACGCAGCAGG	GCAGTCGCCC	TAAAACAAAG	1200
TTAGGTGGCT	CAAGTATGGG	CATCATTCGC	ACATGTAGGC		CCAAGTCAAA	1260
TCCATGCGGG	CTGCTCTTGA	TCTTTTCGGT	CGTGAGTTCG		CACCTACTCC	1320
	CGGACTCCGA	TTACCTCGGG	AACTTGCTCC			1380
CTTGCTGCCT	TCGACCAAGA	AGCGGTTGTT				1440
TTTGAGCAGC	CGCGTAGTGA	GATCTATATC				1500
AGGCAGGGCA	TTGCCACCGC	GCTCATCAAT				1560
	TCTACGTGCA					1620
AAGTTGGGCA	TACGGGAAGA					1680
CAATTCGTTC						1740
TGCGCTGTAT						1800
TCAGGGGGG						1860 1920
TGGATTGATI						1920
GCTCGCGAAG	-				,	2040
GGCTCTATGC						2100
AGCGGTGTTC						2160
GAACTGCTT	GGCAGGCCGG	G CGAGCAGATT	r TCCCAAGGC	G CTGATCACG1	GCTGTGTGTC	2171
GCGGGCTGC	A G					2111

· •

CTGCAGCCGA	GCATCGATTG	AGCACTTTAC	CCAGCTGCGC	TGGCTGACCA	TTCAGAATGG	60
CCCGCGGCAC	TATCCAATCT	AAATCGATCT	TCGGGCGCCG	CGGGCATCAT	GCCCGCGGCG	120
CTCGCCTCAT	TTCAATCTCT	AACTTGATAA	AAACAGAGCT	GTTCTCCGGT	CTTGGTGGAT	180
CAAGGCCAGT	CGCGGAGAGT	CTCGAAGAGG	AGAGTACAGT	GAACGCCGAG	TCCACATTGC	240
AACCGCAGGC	ATCATCATGC	TCTGCTCAGC	CACGCTACCG	CAGTGTGTCG	ATTGGTCATC	300
CTCCGGTTGA	GGTTACGCAA	GACGCTGGAG	GTATTGTCCG	GATGCGTTCT	CTCGAGGCGC	360
TTCTTCCCTT	CCCGGGTCGA	ATTCTTGAGC	GTCTCGAGCA	TTGGGCTAAG	ACCCGTCCAG	420
AACAAACCTG	CGTTGCTGCC	AGGGCGGCAA	ATGGGGAATG	GCGTCGTATC	AGCTACGCGG	480
AAATGTTCCA	CAACGTCCGC	GCCATCGCAC	AGAGCTTGCT	TCCTTACGGA	CTATCGGCAG	540
AGCGTCCGCT	GCTTATCGTC	TCTGGAAATG	ACCTGGAACA	TCTTCAGCTG	GCATTTGGGG	600
CTATGTATGC	GGGCATTCCC	TATTGCCCGG	TGTCTCCTGC	TTATTCACTG	CTGTCGCAAG	660
ATTTGGCGAA	GCTGCGTCAC	ATCGTAGGTC	TTCTGCAACC	GGGACTGGTC	TTTGCTGCCG	720
ATGCAGCACC	TTTCCAGCGC	GCTGTTTTGC	AATGGCGGTC	GGCGAAAGTT	GATGCGCTGT	780
ATCGTGGTGA	AGATCAATCC	ATGCTGCGTG	ACGAGGCCAC	ACTGTGAGTT	GGTCAGGGGG	840
GGCTTACTCG	GCGTTTTCCG	ACACTGCGTT	GGTTGCGGCA	GTGCGCACCC	CCTGGATTGA	900
TTGCGGGGGT	GCCCTGTCGC	TGGTGTCGCC	TATCGACTTA	GGGGTAAAGG	TCGCTCGCGA	960
AGTTCTGATG	CGTGCGTCGC	TTGAACCACA	AATGGTCGAT	AGCGTACTCG	CAGGCTCTAT	1020
GGCTCAAGCA	AGCTTTGATG	CTTACCTGCT	CCCGCGGCAC	ATTGGCTTGT	ACAGCGGTGT	1080
TCCCAAGTCG	GTTCCGGCCT	TGGGGGTGCA	GCGCATTTGC	GGCACAGGCT	TCGAACTGCT	1140
TCGGCAGGCC	GGCGAGCAGA	TTTCCCAAGG	CGCTGATCAC	GTGCTGTGTG	TCGCGGGCTG	1200
CAG						1203

•
•
3.22
•

GAATTCCCCT	GGCGACGAAA	000000	CCGCATGGCC	ACGGCTGGGC	GGTAACTGAT	60
GCTTGCGTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA		AGAATCATGC	120
GGGTACGCCT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC		AAAAATAGTT	180
AATTGACAGA	ACTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG	240
GGTGCACGAT	GAATAGCTAC	GATGGCCGTT	GGTCTACCGT	TGATGTGAAG	GTTGAAGAAG	300
GTATCGCTTG	GGTCACGCTG	AACCGCCCGG	AGAAGCGCAA	CGCAATGAGC	CCAACTCTCA	360
ATCGAGAGAT	GGTCGAGGTT	CTGGAGGTGC	TGGAGCAGGA	CGCAGATGCT	CGCGTGCTTG	420
TTCTGACTGG	TGCAGGCGAA	TCCTGGACCG	CGGGCATGGA	CCTGAAGGAG	TATTTCCGCG	480
AGACCGATGC	TGGCCCCGAA	ATTCTGCAAG	AGAAGATTCG	TCGGGGACAG	CAAGCGAACC	540
GGAATTGCCA	GCTGGGGCGC	CCTCTGGTAA	GGTTGGGAAG	CCCTGCAAAG	TAAACTGGAT	600
GGCTTTCTTG	CCGCCAAGGA	TCTGATGGCG	CAGGGGATCA	AGATCTGATC	AAGAGACAGG	660
ATGAGGATCG	TTTCGCATGA	TTGAACAAGA	TGGATTGCAC	GCAGGTTCTC	CGGCCGCTTG	720
GGTGGAGAGG	CTATTCGGCT	ATGACTGGGC	ACAACAGACA	ATCGGCTGCT	CTGATGCCGC	780
CGTGTTCCGG	CTGTCAGCGC	AGGGGCGCCC	GGTTCTTTTT	GTCAAGACCG	ACCTGTCCGG	840
TGCCCTGAAT	GAACTGCAGG	ACGAGGCAGC	GCGGCTATCG	TGGCTGGCCA	CGACGGGCGT	900
TCCTTGCGCA	GCTGTGCTCG	ACGTTGTCAC	TGAAGCGGGA	AGGGACTGGC	TGCTATTGGG	960
CGAAGTGCCG	GGGCAGGATC	TCCTGTCATC	TCACCTTGCT	CCTGCCGAGA	AAGTATCCAT	1020
CATGGCTGAT	GCAATGCGGC	GGCTGCATAC	GCTTGATCCG	GCTACCTGCC	CATTCGACCA	1080
CCAAGCGAAA	CATCGCATCG	AGCGAGCACG	TACTCGGATG	GAAGCCGGTC	TTGTCGATCA	1140
GGATGATCTG	GACGAAGAGC	ATCAGGGGCT	CGCGCCAGCC	GAACTGTTCG	CCAGGCTCAA	1200
GGCGCGCATG	CCCGACGGCG	AGGATCTCGT	CGTGACCCAT	GGCGATGCCT	GCTTGCCGAA	1260
TATCATGGTG	GAAAATGGCC	GCTTTTCTGG	ATTCATCGAC	TGTGGCCGGC	TGGGTGTGGC	1320
GGACCGCTAT	CAGGACATAG	CGTTGGCTAC	CCGTGATATT	GCTGAAGAGC	TTGGCGGCGA	1380
ATGGGCTGAC	CGCTTCCTCG	TGCTTTACGG	TATCGCCGCT	CCCGATTCGC	AGCGCATCGC	1440
CTTCTATCGC	CTTCTTGACG	AGTTCTTCTG	AGCGGGACTC	TGGGGTTCGA	AATGACCGAC	1500
CAAGCGACGC	CCCGAGCAGG	GCATGAAGCA	GTTCCTTGAC	GAGAAAAGCA	TCAAGCCGGG	1560
CTTGCAGACC	TACAAGCGCT	GATAAATGCG	CCGGGGCCCT	CGCTGCGCCC	CCGGCCTTCC	1620
AATAATGACA	ATAATGAGGA	GTGCCCAATG	TTTCACGTGC	CCCTGCTTAT	TGGTGGTAAG	1680
CCTTGTTCAG	CATCTGATGA	GCGCACCTTC	GAGCGTCGTA	GCCCGCTGAC	CGGAGAAGTG	1740
GTATCGCGCG	TCGCTGCTGC	CAGTTTGGAA	GATGCGGACG	CCGCAGTGGC	CGCTGCACAG	1800
GCTGCGTTTC	CTGAATGGGC	GGCGCTTGCT	CCGAGCGAAC	GCCGTGCCCG	ACTGCTGCGA	1860
GCGGCGGATC	TTCTAGAGGA	CCGTTCTTCC	GAGTTCACCG	CCGCAGCGAG	TGAAACTGGC	1920
GCAGCGGGAA	ACTGGTATGG	GTTTAACGTT	TACCTGGCGG	CGGGCATGTT	GCGGGGAATT	1980
С						1981

			•
			•

GAATTCCCCT	GGCGACGAAA	GGGCGGCAGG	CCGCATGGCC	ACGGCTGGGC	GGTAACTGAT	60
GCTTGCGTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA	TTTCGGCGAG	AGAATCATGC	120
GGGTACGCCT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC	CTTGAATCAG	AAAAATAGTT	180
AATTGACAGA	ACTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG	240
GGTGCACGAT	GAATAGCTAC	GATGGCCGTT	GGTCTACCGT	TGATGTGAAG	GTTGAAGAAG	300
GTATCGCTTG	GGTCACGCTG	AACCGCCCGG	AGAAGCGCAA	CGCAATGAGC	CCAACTCTCA	360
ATCGAGAGAT	GGTCGAGGTT	CTGGAGGTGC	TGGAGCAGGA	CGCAGATGCT	CGCGTGCTTG	420
TTCTGACTGG	TGCAGGCGAA		CGGGCATGGA		TATTTCCGCG	480
AGACCGATGC				TCGGGGGAGA	GGCGGTTTGC	540
GTATTGGGCG	CATGCATAAA	AACTGTTGTA	ATTCATTAAG	CATTCTGCCG	ACATGGAAGC	600
CATCACAAAC	GGCATGATGA		CCAGCGGCAT	CAGCACCTTG	TCGCCTTGCG	660
TTTATAATAT	GCCCATGGAC					720
ATAAGCCTGT				ATGCGCTCAC		780
CAGAACCTTG	ACCGAACGCA					840
TGACTGTTTT	TTTGTACAGT	CTATGCCTCG	GGCATCCAAG	CAGCAAGCGC	GTTACGCCGT	900
GGGTCGATGT		GGAGCAGCAA		CAGCAGCAAC		960
AGCAGGGCAG	TCGCCCTAAA	ACAAAGTTAG	GTGGCTCAAG	TATGGGCATC	ATTCGCACAT	1020
GTAGGCTCGG	CCCTGACCAA	GTCAAATCCA	TGCGGGCTGC	TCTTGATCTT	TTCGGTCGTG	1080
AGTTCGGAGA	CGTAGCCACC			CTCCGATTAC	CTCGGGAACT	1140
TGCTCCGTAG	TAAGACATTC	ATCGCGCTTG	CTGCCTTCGA	CCAAGAAGCG	GTTGTTGGCG	1200
CTCTCGCGGC	TTACGTTCTG	• •	AGCAGCCGCG			1260
ATCTCGCAGT	CTCCGGCGAG	•		CACCGCGCTC		1320
TCAAGCATGA	GGCCAACGCG	CTTGGTGCTT		CGTGCAAGCA		1380
ACGATCCCGC	AGTGGCTCTC	TATACAAAGT	TGGGCATACG	GGAAGAAGTG	ATGCACTTTG	1440
ATATCGACCC	AAGTACCGCC	ACCTAACAAT		CGAGATCGGC	TTCCCCGAGC	1500
AGGGCATGAA	GCAGTTCCTT	GACGAGAAAA	GCATCAAGCC			1560
GCTGATAAAT	GCGCCGGGGC	CCTCGCTGCG			ACAATAATGA	1620
GGAGTGCCCA	ATGTTTCACG	TGCCCCTGCT		AAGCCTTGTT		1680
TGAGCGCACC	TTCGAGCGTC	GTAGCCCGCT		GTGGTATCGC		1740
TGCCAGTTTG	GAAGATGCGG	ACGCCGCAGT		CAGGCTGCGT	TTCCTGAATG	1800
GGCGGCGCTT	GCTCCGAGCG			• •	ATCTTCTAGA	1860
GGACCGTTCT	TCCGAGTTCA			GGCGCAGCGG	GAAACTGGTA	1920
TGGGTTTAAC	GTTTACCTGG	CGGCGGGCAT	GTTGCGGGGA	ATTC		1964

			•
			•
			•

GAATTCCCCT	GGCGACGAAA	GGGCGGCAGG	CCGCATGGCC	ACGGCTGGGC	GGTAACTGAT	60
GCTTGCGTTA	ATCGTTAACC	GTTTGAAATT	CCTTGCCAAA	TTTCGGCGAG	AGAATCATGC	120
GGGTACGCCT	TTCCGTGCGC	TTTGATCTGC	GCTTCCGTGC	CTTGAATCAG	AAAAATAGTT	180
AATTGACAGA	ACTATAGGTT	CGCAGTAGCT	TTTGCTCACC	CACCAAATCC	ACAGCACTGG	240
GGTGCACGAT	GAATAGCTAC	GATGGCCGTT	GGTCTACCGT	TGATGTGAAG	GTTGAAGAAG	300
GTATCGCTTG	GGTCACGCTG	AACCGCCCGG	AGAAGCGCAA	CGCAATGAGC	CCAACTCTCA	360
ATCGAGAGAT	GGTCGAGGTT	CTGGAGGTGC	TGGAGCAGGA	CGCAGATGCT	CGCGTGCTTG	420
TTCTGACTGG	TGCAGGCGAA	TCCTGGACCG	CGGGCATGGA	CCTGAAGGAG	TATTTCCGCG	480
AGACCGATGC	TGGCCCCGAA	ATTCTGCAAG	AGAAGATTCG	TCGCGAGCAG	GGCATGAAGC	540
AGTTCCTTGA	CGAGAAAAGC	ATCAAGCCGG	GCTTGCAGAC	CTACAAGCGC	TGATAAATGC	600
GCCGGGGCCC	TCGCTGCGCC	CCCGGCCTTC	CAATAATGAC	AATAATGAGG	AGTGCCCAAT	660
GTTTCACGTG		TTGGTGGTAA	GCCTTGTTCA	GCATCTGATG	AGCGCACCTT	720
CGAGCGTCGT	AGCCCGCTGA	CCGGAGAAGT	GGTATCGCGC	GTCGCTGCTG	CCAGTTTGGA	780
AGATGCGGAC		CCGCTGCACA	GGCTGCGTTT	CCTGAATGGG	CGGCGCTTGC	840
TCCGAGCGAA	CGCCGTGCCC	GACTGCTGCG	AGCGGCGGAT	CTTCTAGAGG	ACCGTTCTTC	900
CGAGTTCACC	GCCGCAGCGA	GTGAAACTGG	CGCAGCGGGA	AACTGGTATG	GGTTTAACGT	960
TTACCTGGCG	GCGGGCATGT	TGCGGGGAAT	TC			992

		•
		•
		•
		•

	AATGACAATA					60
	TGTTCAGCAT					120
	TCGCGCGTCG					180
TGCACAGGCT	GCGTTTCCTG	AATGGGCGGC	GCTTGCTCCG	AGCGAACGCC	GTGCCCGACT	240
GCTGCGAGCG	GCGGATCTTC	TAGAGGACCG	TTCTTCCGAG	TTCACCGCCG	CAGCGAGTGA	300
AACTGGCGCA	GCGGGAAACT	GGTATGGGTT		CTGGCGGCGG		360
GGAAGCCGCG	GCCATGACCA					420
TAGCTTTGCC	ATGGCGGTTC	GACAGCCATG	TGGCGTGGTG	CTCGGTATTG	CGCCTTGGAA	480
TGCTCCGGTA	ATCCTTGGCG	TACGGGCTGT	TGCGATGCCG	TTGGCATGCG	GCAATACCGT	540
GGTGTTGAAA	AGCTCTGAGC	TGAGTCCCTT	TACCCATCGC	CTGATTGGTC	AGGTGTTGCA	600
TGATGCTGGT	CTGGGGGATG	GCGTGGTGAA	TGTCATCAGC	AATGCCCCGC	AAGACGCTCC	660
TGCGGTGGTG	GAGCGACTGA	TTGCAAATCC	TGCGGTACGT	CGAGTGAACT	TCACCGGTTC	720
	GGACGGATCA			CATCTGAAGC		780
	GGTAAGGCTC					840
CGAAGCGGCG	GCCTTTGGTG	CCTACTTCAA	TCAGGGTCAA	ATCTGCATGT	CCACTGAGCG	900
	ACAGCAGTCG			CTGGCGAGGA		960
ACTGCGTGCT	GGCGATCCTA					1020
TGCAGGTCAA	CGCATCCAGG			GGGGACAGCA		1080
AATTGCCAGC	TGGGGCGCCC	TCTGGTAAGG	TTGGGAAGCC	CTGCAAAGTA	AACTGGATGG	1140
	GCCAAGGATC			ATCTGATCAA		1200
	TCGCATGATT					1260
	ATTCGGCTAT					1320
	GTCAGCGCAG					1380
	ACTGCAGGAC					1440
CTTGCGCAGC	TGTGCTCGAC	GTTGTCACTG	AAGCGGGAAG	GGACTGGCTG	CTATTGGGCG	1500
	GCAGGATCTC					1560
TGGCTGATGC	AATGCGGCGG				TTCGACCACC	1620
AAGCGAAACA	TCGCATCGAG				GTCGATCAGG	1680
ATGATCTGGA	CGAAGAGCAT				AGGCTCAAGG	1740
CGCGCATGCC	CGACGGCGAG	GATCTCGTCG	TGACCCATGG	CGATGCCTGC		1800
TCATGGTGGA	AAATGGCCGC		TCATCGACTG			1860
ACCGCTATCA	GGACATAGCG			TGAAGAGCTT		1920
GGGCTGACCG	CTTCCTCGTG				CGCATCGCCT	1980
TCTATCGCCT	TCTTGACGAG				TGACCGACCA	2040
	GGCCCAGCGC				CCGACTGTGC	2100
ATGACGAGGC	TCAGATGCCA	TTCGGTGGGG	TGAAGTCCAG	CGGCTACGGC	AGCTTCGGCA	2160

		*

GTCGAGCATC	GATTGAGCAC	TTTACCCAGC	TGCGCTGGCT	GACCATTCAG	AATGGCCCGC	2220
GGCACTATCC	AATCTAAATC	GATCTTCGGG	CGCCGCGGC	ATCATGCCCG	CGGCGCTCGC	2280
CTCATTTCAA	TCTCTAACTT	GATAAAAACA	GAGCTGTTCT	CCGGTCTTGG	TGGATCAAGG	2340
CCAGTCGCGG	AGAGTCTCGA	AGAGGAGAGT	ACAGTGAACG	CCGAGTCCAC	ATTGCAACCG	2400
CAGGCATCAT	CATGCTCTGC	TCAGCCACGC	TACCGCAGTG	TGTCGATTGG	TCATCCTCCG	2460
GTTGAGGTTA	CGCAAGACGC	TGGAGGTATT	GTCCGGATGC	GTTCTCTCGA	GGCGCTTCTT	2520
CCCTTCCCGG	GTGGAATTC					2539

		•
		•
		+

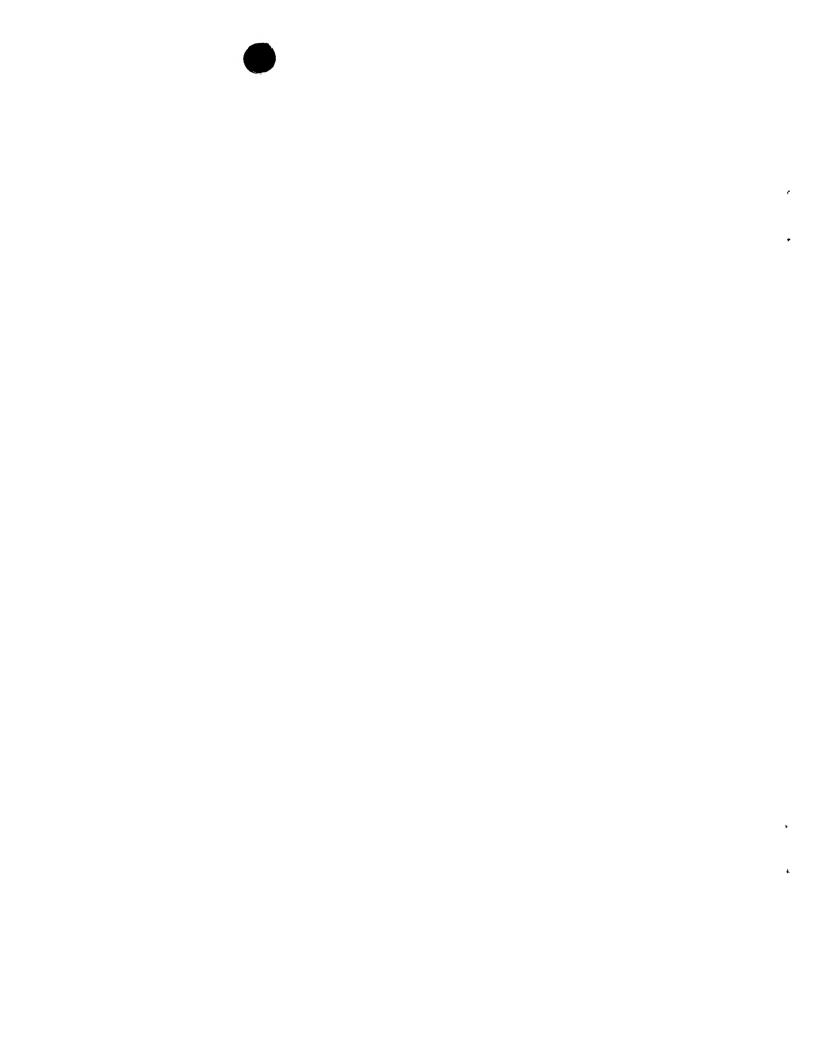
GAATTCCAAT						60
				CGTCGTAGCC		120
				GCGGACGCCG		180
				AGCGAACGCC		240
				TTCACCGCCG		300
				CTGGCGGCGG		360
				ATTCCGTCCA		42,0
TAGCTTTGCC	ATGGCGGTTC	GACAGCCATG	TGGCGTGGTG	CTCGGTATTG	CGCCTTGGAA	480
		TACGGGCTGT		TTGGCATGCG		540
				CTGATTGGTC		600
				AATGCCCCGC		660
				CGAGTGAACT		720
				CATCTGAAGC		780
				GCCGACCTCG		840
				ATCTGCATGT		900
				CTGGCGAGGA		960
				GGTTCGTTGA		1020
				ATTGGGCGCA		1080
				TCACAAACGG		1140
				TAATATTTGC		1200
				AAGCCTGTTC		1260
				GAACCTTGAC		1320
				ACTGTTTTTT		1380
				GTCGATGTTT		1440
				CAGGGCAGTC		1500
				AGGCTCGGCC		1560
				TTCGGAGACG		1620
				CTCCGTAGTA		1680
				CTCGCGGCTT		1740
				CTCGCAGTCT		1800
				AAGCATGAGG		1860
				GATCCCGCAG		1920
				ATCGACCCAA		1980
				CCAGCGCGTC		2040
				GATGCCATTC		2100
AGTCCAGCGG	CTACGGCAGC	TTCGGCAGTC	GAGCATCGAT	TGAGCACTTT	ACCCAGCTGC	2160

		,
		•
		•

GCTGGCTGAC	CATTCAGAAT	GGCCCGCGGC	ACTATCCAAT	CTAAATCGAT	CTTCGGGCGC	2220
					AAAAACAGAG	
					GGAGAGTACA	
0101					GCCACGCTAC	
0.0						
					AGGTATTGTC	
CGGATGCGTT	CTCTCGAGGC	GCTTCTTCCC	TTCCCGGGTG	GAATTC		2506

		•

GAATTCCAAT	AATGACAATA	ATGAGGAGTG	CCCAATGTTT	CACGTGCCCC	TGCTTATTGG	60
TGGTAAGCCT	TGTTCAGCAT	CTGATGAGCG	CACCTTCGAG	CGTCGTAGCC	CGCTGACCGG	120
AGAAGTGGTA	TCGCGCGTCG	CTGCTGCCAG	TTTGGAAGAT	GCGGACGCCG	CAGTGGCCGC	180
TGCACAGGCT	GCGTTTCCTG	AATGGGCGGC	GCTTGCTCCG	AGCGAACGCC	GTGCCCGACT	240
GCTGCGAGCG	GCGGATCTTC	TAGAGGACCG	TTCTTCCGAG	TTCACCGCCG	CAGCGAGTGA	300
AACTGGCGCA	GCGGGAAACT	GGTATGGGTT	TAACGTTTAC	CTGGCGGCGG	GCATGTTGCG	360
GGAAGCCGCG	GCCATGACCA	CACAGATTCA	GGGCGATGTC	ATTCCGTCCA	ATGTGCCCGG	420
TAGCTTTGCC	ATGGCGGTTC	GACAGCCATG	TGGCGTGGTG	CTCGGTATTG	CGCCTTGGAA	480
TGCTCCGGTA	ATCCTTGGCG	TACGGGCTGT	TGCGATGCCG	TTGGCATGCG	GCAATACCGT	540
GGTGTTGAAA	AGCTCTGAGC	TGAGTCCCTT	TACCCATCGC	CTGATTGGTC	AGGTGTTGCA	600
TGATGCTGGT	CTGGGGGATG	GCGTGGTGAA	TGTCATCAGC	AATGCCCCGC	AAGACGCTCC	660
TGCGGTGGTG	GAGCGACTGA	TTGCAAATCC	TGCGGTACGT	CGAGTGAACT	TCACCGGTTC	720
GACCCACGTT	GGACGGATCA	TTGGTGAGCT	GTCTGCGCGT	CATCTGAAGC	CTGCTGTGCT	780
GGAATTAGGT	GGTAAGGCTC	CGTTCTTGGT	CTTGGACGAT	GCCGACCTCG	ATGCGGCGGT	840
CGAAGCGGCG	GCCTTTGGTG	CCTACTTCAA	TCAGGGTCAA	ATCTGCATGT	CCACTGAGCG	900
TCTGATTGTG	ACAGCAGTCG	CAGACGCCTT	TGTTGAAAAG	CTGGCGAGGA	AGGTCGCCAC	960
ACTGCGTGCT	GGCGATCCTA	ATGATCCGCA	ATCGGTCTTG	GGTTCGTTGA	TTGATGCCAA	1020
TGCAGGTCAA	CGCATCCAGG	TTCTGGTCGA	TGATGCGCTC	GCAAAAGGCG	CGCAATGGAA	1080
TTGGCCCAGC	GCGTCGATTC	GGGCATTTGC	CATATCAATG	GACCGACTGT	GCATGACGAG	1140
GCTCAGATGC	CATTCGGTGG	GGTGAAGTCC	AGCGGCTACG	GCAGCTTCGG	CAGTCGAGCA	1200
TCGATTGAGC	ACTTTACCCA	GCTGCGCTGG	CTGACCATTC	AGAATGGCCC	GCGGCACTAT	1260
CCAATCTAAA	TCGATCTTCG	GGCGCCGCGG	GCATCATGCC	CGCGGCGCTC	GCCTCATTTC	1320
AATCTCTAAC	TTGATAAAAA	CAGAGCTGTT	CTCCGGTCTT	GGTGGATCAA	GGCCAGTCGC	1380
GGAGAGTCTC	GAAGAGGAGA	GTACAGTGAA	CGCCGAGTCC		CGCAGGCATC	1440
ATCATGCTCT	GCTCAGCCAC		TGTGTCGATT		CGGTTGAGGT	1500
TACGCAAGAC	GCTGGAGGTA	TTGTCCGGAT	GCGTTCTCTC	GAGGCGCTTC	TTCCCTTCCC	1560
GGGTGGAATT	С					1571



GAATTCCGCG	GTCGGCGAAA	GTTGATGCGC	TGTATCGTGG	TGAAGATCAA	TCCATGCTGC	60
GTGACGAGGC	CACACTGTGA	GTTGGTCAGG	GGGGGCTTAC	TCGGCGTTTT	CCGACACTGC	120
GTTGGTTGCG	GCAGTGCGCA	CCCCTGGAT	TGATTGCGGG	GGTGCCCTGT	CGCTGGTGTC	180
GCCTATCGAC	TTAGGGGTAA	AGGTCGCTCG	CGAAGTTCTG	ATGCGTGCGT	CGCTTGAACC	240
ACAAATGGTC	GATAGCGTAC	TCGCAGGCTC	TATGGCTCAA	GCAAGCTTTG	ATGCTTACCT	300
GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	TGTTCCCAAG	TCGGTTCCGG	CCTTGGGGGT	360
GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	GCTTCGGCAG	GCCGGCGAGC	AGATTTCCCA	42.0
AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGC	AGAGTCCATG	TCGCGTAACC	CCATCGCGTC	480
				GAGTTCAAGG		540
GGAGGCATTG	TTTGATCCTG	CTCCAGGACT	CGACATGATC	GCTACCGCAG	AAAACCTGGG	600
GACAGCAAGC	GAACCGGAAT	TGCCAGCTGG	GGCGCCCTCT	GGTAAGGTTG	GGAAGCCCTG	660
				TGGCGCAGGG		720
TGATCAAGAG	ACAGGATGAG	GATCGTTTCG	CATGATTGAA	CAAGATGGAT	TGCACGCAGG	780
TTCTCCGGCC	GCTTGGGTGG	AGAGGCTATT	CGGCTATGAC	TGGGCACAAC	AGACAATCGG	840
CTGCTCTGAT	GCCGCCGTGT	TCCGGCTGTC	AGCGCAGGGG	CGCCCGGTTC	TTTTTGTCAA	900
GACCGACCTG	TCCGGTGCCC	TGAATGAACT	GCAGGACGAG	GCAGCGCGGC	TATCGTGGCT	960
GGCCACGACG	GGCGTTCCTT	GCGCAGCTGT	GCTCGACGTT	GTCACTGAAG	CGGGAAGGGA	1020
CTGGCTGCTA	TTGGGCGAAG	TGCCGGGGCA	GGATCTCCTG	TCATCTCACC	TTGCTCCTGC	1080
CGAGAAAGTA	TCCATCATGG	CTGATGCAAT	GCGGCGGCTG	CATACGCTTG	ATCCGGCTAC	1140
CTGCCCATTC	GACCACCAAG	CGAAACATCG	CATCGAGCGA	GCACGTACTC	GGATGGAAGC	1200
CGGTCTTGTC	GATCAGGATG	ATCTGGACGA	AGAGCATCAG	GGGCTCGCGC	CAGCCGAACT	1260
GTTCGCCAGG	CTCAAGGCGC	GCATGCCCGA	CGGCGAGGAT	CTCGTCGTGA	CCCATGGCGA	1320
TGCCTGCTTG	CCGAATATCA	TGGTGGAAAA	TGGCCGCTTT	TCTGGATTCA	TCGACTGTGG	1380
CCGGCTGGGT	GTGGCGGACC	GCTATCAGGA	CATAGCGTTG	GCTACCCGTG	ATATTGCTGA	1440
AGAGCTTGGC	GGCGAATGGG	CTGACCGCTT	CCTCGTGCTT	TACGGTATCG	CCGCTCCCGA	1500
TTCGCAGCGC	ATCGCCTTCT	ATCGCCTTCT	TGACGAGTTC	TTCTGAGCGG	GACTCTGGGG	1560
TTCGAAATGA	CCGACCAAGC	GACGCCCATT	GAGGGCGCAA	GAGGAGAAAT	GGATTGACCA	1620
AGAGATCGTG	GCTGTTACGG	ATGAACAGTT	CGATTTAGAG	GGCTACAACA	GTCGAGCAAT	1680
TGAACTGCCT	CGGAAGGCAA	AATTGTTGAT	CGTGACAGTC	ATCCGCGGCC	TAGCAGTCTT	1740
TGAAGCCCTT	TCCCGATTGA	AGCCTGTTCA	TTCTGGCGGG	GTGCAGACTG	CGGGCAACAG	1800
CTGTGCCGTA	GTGGACGGCG	CCGCGGCGGC	TTTGGTGGCT	CGAGAGTCGT	CTGCGACACA	1860
GCCGGTCTTG	GCTAGGATAC	TGGCTACCTC	CGTAGTCGGG	ATCGAGCCCG	AGCATATGGG	1920
GCTCGGCCCT	GCGCCCGCGA	TTCGCCTGCT	GCTTGCGCGT	' AGTGATCTTA	GTTTGAGGGA	1980
TATCGACCTC					TACAGCATGA	2040
ATTGGGTATT	GAGCACTCAA	AACTTAATAT	TTGGGGCGGG	GCCATTGCAC	TTGGACACCC	2100
GCTTGCCGCG	ACCGGATTGC	GTCTCTGCAT	GACCCTCGCT	CACCAATTGC	AAGCTAATAA	2160

•
•
·
4

CTTTCGATAT GGAATTGCCT CGGCATGCAT TGGTGGGGGA CAGGGGATGG CGGTTCTTTT 2220
AGAGAATCCC CACTTCGGTT CGTCCTCTCC ACGAAGTTCG ATGATTAACA GAGTTGACCA 2280
CTATCCACTG AGCTAACGGG CATCTCCTTT GTTGCTTTGA GGTGGCGCAC GAAGGAGGC 2340
TCGAAAATCT CTGCTAAAAA CAAGAAGAAG GAACAGGGAA CATGATTAGT TTCGCTCGTA 2400
TCGGGCTGAT TGTTACCGGC ACGGCTTACC TTGCCCTTGC CTTCGCACTC GTATTATGTG 2460
TCGGGCTGAT TGTTACCGGC ACGGGTTTCT ACAGTGTACA TACCTTGTCA GGGTTGGTGG 2520
GAATTC 2220

	•
	•
	U
	4

GAATTCCGCG	GTCGGCGAAA	GTTGATGCGC	TGTATCGTGG	TGAAGATCAA	TCCATGCTGC	60
	CACACTGTGA					120
GTTGGTTGCG	GCAGTGCGCA	CCCCCTGGAT	TGATTGCGGG	GGTGCCCTGT	CGCTGGTGTC	180
	TTAGGGGTAA					240
	GATAGCGTAC			GCAAGCTTTG	ATGCTTACCT	300
GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	TGTTCCCAAG	TCGGTTCCGG		360
GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	GCTTCGGCAG	GCCGGCGAGC	AGATTTCCCA	420
AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGC	AGAGTCCATG	TCGCGTAACC	CCATCGCGTC	480
GTATACACAC	CGGGGCGGGT			GAGTTCAAGG	ATTTTTTGTG	540
GGAGGCATTG	TTTGATCCTG	CTCCAGGACT	CGACATGATC	GCTACCGCAG	AAAACCTGGG	600
GGAGAGGCGG	TTTGCGTATT			TTGTAATTCA		660
TGCCGACATG				AATCGCCAGC		720
CCTTGTCGCC				CCGTGGAAAC		780
ACGAACCCAG	TTGACATAAG			TAATGCAAGT		840
	TGGTCCAGAA					900
	TGTTATGACT					960
AGCGCGTTAC	GCCGTGGGTC	GATGTTTGAT	GTTATGGAGC	AGCAACGATG	TTACGCAGCA	1020
	TACGCAGCAG					1080
GCATCATTCG	CACATGTAGG	CTCGGCCCTG	ACCAAGTCAA	ATCCATGCGG	GCTGCTCTTG	1140
ATCTTTTCGG	TCGTGAGTTC	GGAGACGTAG	CCACCTACTC	CCAACATCAG	CCGGACTCCG	1200
ATTACCTCGG	GAACTTGCTC	CGTAGTAAGA	CATTCATCGC	GCTTGCTGCC	TTCGACCAAG	1260
AAGCGGTTGT				GTTTGAGCAG		1320
AGATCTATAT	CTATGATCTC	GCAGTCTCCG	GCGAGCACCG	GAGGCAGGGC	ATTGCCACCG	1380
CGCTCATCAA	TCTCCTCAAG	CATGAGGCCA	ACGCGCTTGG	TGCTTATGTG	ATCTACGTGC	1440
AAGCAGATTA	CGGTGACGAT	CCCGCAGTGG	CTCTCTATAC	AAAGTTGGGC	ATACGGGAAG	1500
AAGTGATGCA	CTTTGATATC	GACCCAAGTA	CCGCCACCTA	ACAATTCGTT	CAAGCCGAGA	1560
	ATTGAGGGCG					1620
CGGATGAACA	GTTCGATTTA	GAGGGCTACA	ACAGTCGAGC	AATTGAACTG	CCTCGGAAGG	1680
CAAAATTGTT	GATCGTGACA	GTCATCCGCG	GCCTAGCAGT	CTTTGAAGCC	CTTTCCCGAT	1740
TGAAGCCTGT					GTAGTGGACG	1800
	GGCTTTGGTG					1860
	CTCCGTAGTC					1920
					CTCTTTGAGA	1980
					ATTGAGCACT	2040
					GCGACCGGAT	2100
TGCGTCTCTG	CATGACCCTC	GCTCACCAAT	TGCAAGCTAA	. TAACTTTCGA	TATGGAATTG	2160

		ŕ
		•
		•

WO 00/26355 PCT/EP99/07952

CCTCGGCATG	CATTGGTGGG	GGACAGGGGA	TGGCGGTTCT	TTTAGAGAAT	CCCCACTTCG	2220
GTTCGTCCTC	TGCACGAAGT	TCGATGATTA	ACAGAGTTGA	CCACTATCCA	CTGAGCTAAC	2280
GGGCATCTCC	TTTGTTGCTT	TGAGGTGGCG	CACGAAGGAG	GGCTCGAAAA	TCTCTGCTAA	2340
AAACAAGAAG	AAGGAACAGG	GAACATGATT	AGTTTCGCTC	GTATGGCAGA	AAGTTTAGGA	2400
GTCCAGGCTA	AACTTGCCCT	TGCCTTCGCA	CTCGTATTAT	GTGTCGGGCT	GATTGTTACC	2460
GGCACGGGTT	TCTACAGTGT	ACATACCTTG	TCAGGGTTGG	TGGGAATTC		2509

1

		•
		•
		·
		J

GAATTCCGCG	GTCGGCGAAA	GTTGATGCGC	TGTATCGTGG	TGAAGATCAA	TCCATGCTGC	60
GTGACGAGGC	CACACTGTGA	GTTGGTCAGG	GGGGGCTTAC	TCGGCGTTTT	CCGACACTGC	120
GTTGGTTGCG	GCAGTGCGCA	CCCCTGGAT	TGATTGCGGG	GGTGCCCTGT	CGCTGGTGTC	180
GCCTATCGAC	TTAGGGGTAA	AGGTCGCTCG	CGAAGTTCTG	ATGCGTGCGT	CGCTTGAACC	240
ACAAATGGTC	GATAGCGTAC	TCGCAGGCTC	TATGGCTCAA	GCAAGCTTTG	ATGCTTACCT	300
GCTCCCGCGG	CACATTGGCT	TGTACAGCGG	TGTTCCCAAG	TCGGTTCCGG	CCTTGGGGGT	360
GCAGCGCATT	TGCGGCACAG	GCTTCGAACT	GCTTCGGCAG	GCCGGCGAGC	AGATTTCCCA	420
AGGCGCTGAT	CACGTGCTGT	GTGTCGCGGC	AGAGTCCATG	TCGCGTAACC	CCATCGCGTC	480
GTATACACAC	CGGGGCGGGT	TCCGCCTCGG	TGCGCCCGTT	GAGTTCAAGG	ATTTTTTGTG	540
GGAGGCATTG	TTTGATCCTG	CTCCAGGACT	CGACATGATC	GCTACCGCAG	AAAACCTGGC	600
GCGCATTGAG	GGCGCAAGAG	GAGAAATGGA	TTGACCAAGA	GATCGTGGCT	GTTACGGATG	660
AACAGTTCGA	TTTAGAGGGC	TACAACAGTC	GAGCAATTGA	ACTGCCTCGG	AAGGCAAAAT	720
TGTTGATCGT	GACAGTCATC	CGCGGCCTAG	CAGTCTTTGA	AGCCCTTTCC	CGATTGAAGC	780
CTGTTCATTC	TGGCGGGGTG	CAGACTGCGG	GCAACAGCTG	TGCCGTAGTG	GACGGCGCCG	840
CGGCGGCTTT	GGTGGCTCGA	GAGTCGTCTG	CGACACAGCC	GGTCTTGGCT	AGGATACTGG	900
CTACCTCCGT	AGTCGGGATC	GAGCCCGAGC	ATATGGGGCT	CGGCCCTGCG	CCCGCGATTC	960
GCCTGCTGCT	TGCGCGTAGT	GATCTTAGTT	TGAGGGATAT	CGACCTCTTT	GAGATAAACG	1020
AGGCGCAGGC	CGCCCAAGTT	CTAGCGGTAC	AGCATGAATT	GGGTATTGAG	CACTCAAAAC	1080
TTAATATTTG	GGGCGGGCC	ATTGCACTTG	GACACCCGCT	TGCCGCGACC	GGATTGCGTC	1140
TCTGCATGAC	CCTCGCTCAC	CAATTGCAAG	CTAATAACTT	TCGATATGGA	ATTGCCTCGG	1200
CATGCATTGG	TGGGGGACAG	GGGATGGCGG	TTCTTTTAGA	GAATCCCCAC	TTCGGTTCGT	1260
CCTCTGCACG	AAGTTCGATG	ATTAACAGAG	TTGACCACTA	TCCACTGAGC	TAACGGGCAT	1320
CTCCTTTGTT	GCTTTGAGGT	GGCGCACGAA	GGAGGGCTCG	AAAATCTCTG	CTAAAAACAA	1380
GAAGAAGGAA	CAGGGAACAT	GATTAGTTTC	GCTCGTATGG	CAGAAAGTTT	AGGAGTCCAG	1440
GCTAAACTTG	CCCTTGCCTT	CGCACTCGTA	TTATGTGTCG	GGCTGATTGT	TACCGGCACG	1500
GGTTTCTACA	GTGTACATAC	CTTGTCAGGG	TTGGTGGGAA	TTC		1543

